AD-A260 250

AL-TR-1992-0105

LEADERSHIP EFFECTIVENESS ASSESSMENT PROFILE (LEAP): OFFICER INSTRUMENT FIELD TESTING AND REFINEMENT

Victor H. Appel
Carol Murray Quintana
Richard W. Cole
Mark D. Shermis
Paul D. Grubb

UES, Incorporated
4401 Dayton-Xenia Road
Dayton, OH 45332

Thomas W. Watson

HUMAN RESOURCES DIRECTORATE
MANPOWER AND PERSONNEL RESEARCH DIVISION
Brooks Air Force Base, TX 78235-5352

Amie Headley-Goode

Texas A&M University
College Station, TX 77843-4235

DTIC
ELECTE
FEB 17 1993
S E D

December 1992

Final Report for Period January 1990 – February 1992

Approved for public release; distribution is unlimited.

93-03028

AIR FORCE MATERIEL COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5000

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Office of Public Affairs has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This report has been reviewed and is approved for publication.

Thomas W. Watson
THOMAS W. WATSON
Project Scientist

William E. Alley
WILLIAM E. ALLEY, Technical Director
Manpower and Personnel Research Division

Roger W. Alford
ROGER W. ALFORD, Lt Col, USAF
Chief, Manpower and Personnel Research Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December 1992	Final – January 1990 – February 1992

4. TITLE AND SUBTITLE
Leadership Effectiveness Assessment Profile (LEAP): Officer Instrument Field Testing and Refinement

5. FUNDING NUMBERS
C – F41689-86-D-0052
PE – 62205F
PR – 1121
TA – 13
WU – 01

6. AUTHOR(S)
Victor H. Appel
Carol Murray Quintana
Richard W. Cole
Mark D. Shermis
Paul D. Grubb
Thomas W. Watson
Amie Headley-Goode

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
UES, Incorporated
4401 Dayton-Xenia Road
Dayton, OH 45332

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES)
Armstrong Laboratory
Human Resources Directorate
Manpower and Personnel Research Division
Brooks Air Force Base, TX 78235-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AL-TR-1992-0105

11. SUPPLEMENTARY NOTES
Armstrong Laboratory Technical Monitor: Dr. Thomas W. Watson, (210) 536-3876

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This report documents the field testing and refinement of a biographical inventory called the Leadership Effectiveness Assessment Profile (LEAP). The LEAP is designed to measure leadership and management potential, the ability to function well in team situations, propensity for commitment to the Air Force, and related attributes. Several iterations of LEAP field testing and refinement are discussed. In each case officer cadets or junior officers were used as subjects. Early field testing included gaining extensive feedback from subjects for instrument refinement. Thus, they became co-developers of the LEAP. Later field testing, when more subjects and some adequate criteria were available, involved more standard psychometric assessment. Field testing was moderately successful, although further work will be needed to prepare the LEAP for operational use.

14. SUBJECT TERMS
Biodata
Leadership
Personnel selection

15. NUMBER OF PAGES
132

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

CONTENTS

CONTENTS (Continued)

LIST OF TABLES

LIST OF TABLES (Concluded)

LIST OF FIGURES

Accesion For
NTIS CRA&I
DTIC TAB
Unannounced
Justification
By
Distribution /
Availability Codes
Dist | Avail and / or Special
A-1

DTIC QUALITY INSPECTED 3

PREFACE

This research was accomplished under Task 44, Contract Number F41689-86-D-0052, (SB 58160432), Subcontract Number S-788-044-001, *Field Testing/Refinement of the Leadership Effectiveness Assessment Profile (LEAP)*, with UES, Inc. of Dayton, Ohio. The task was subcontracted to Victor H. Appel, Ph.D. and Associates of Austin, Texas. This study was a continuation of initial efforts by the Individual Attributes Branch, Manpower and Personnel Research Division, Human Resources Directorate (AL/HR), of the Armstrong Laboratory, to develop a biographical measure to improve officer selection and classification.

We thank the many people at Lackland, Randolph, Brooks, Keesler and other Air Force bases who participated as respondents, helped arrange for data collection, or who provided administrative support. We gratefully acknowledge the assistance of personnel research scientists including Dr. Malcolm James Ree, Senior Scientist, Force Management Systems Branch; Dr. Lonnie D. Valentine, Jr., Chief, Individual Attributes Branch; Mr. James A. Earles, Mr. Leasley K. Besetsny, and Ms Nancy J. Allin, Individual Attributes Branch. Ms Sandy Buchanan and Ms Laurel Betz provided excellent data base support. We are also indebted to Dr. William C. Howell, formerly Chief Scientist, Armstrong Laboratory, Human Resources Directorate; Colonel Michael W. Birdlebough, formerly Director, Human Resources Directorate; Lt Col Roger W. Alford, Chief, Manpower and Personnel Research Division of AL/HR; and to Dr. William E. Alley, Technical Director, Manpower and Personnel Research Division, for their guidance and support.

We thank a variety of persons who gave administrative or other support, including Major Mary T. Brown. Keesler AFB; Dr. Judy C. Turner, Headquarters, OTS, Lackland AFB; Colonel Charles Walters, Deputy Commander for Military Training, OTS, Lackland AFB; and Mr. Charles Hamilton, Headquarters, Military Personnel Center, Randolph AFB. For indispensable administrative support at AFROTC summer encampments, we are grateful to Major Donald Dorough and 1Lt Chellie Cameron, Headquarters, AFROTC, Maxwell AFB, and to the following persons: Major William Cox, and Lt Penson, AFROTC Field Training, McConnell AFB, Kansas; Lt Col James Geyer, AFROTC/FTU, Vandenberg AFB, California; Captain Robyn King, AFROTC Field Training, Plattsburgh AFB, New York; Colonel Robert Robbins and Major Cain, AFROTC Field Training, Lackland, AFB, Texas; Lt Colonel Gary D. Baxter, AFROTC Field Training, Lowry AFB, Colorado.

We are especially grateful to the four members of the Laboratory Advisory Group (LAG), of the Armstrong Laboratory, who gave independent critiques of the LEAP instrument and project. These members were: Captain Shirley M. Ross, Headquarters, USAF, The Pentagon; Dr. Michael D. Mumford, George Mason University; Dr. Lawrence H. Peters, Texas Christian University; and Dr. Fred A. Mael, Army Research Institute.

We thank a number of persons directly related to the LEAP project: Dr. Eric Elder and Mr. Brad Falkenberg, LEAP Research Associates, for their dedicated and extended work on behalf of the project; Dr. C. Paul Sparks, formerly of the University of Houston, and Dr. William A. Owens, of the University of Georgia, for their considerable assistance as biodata consultants; Dr. William A. Hays, The University of Texas at Austin, for his advice on psychometric issues. We also thank Dr. Gary Hanson, The University of Texas at Austin, for his continuing conceptual, statistical and editorial assistance to the project.

We also thank Ms Cindy Ehlert and Ms Virginia Wilson who provided extensive word processing and technical editing support just prior to printing. Ms Nancy J. Allin was also very helpful in giving the manuscript a final review.

LEADERSHIP EFFECTIVENESS ASSESSMENT PROFILE (LEAP): FIELD TESTING AND REFINEMENT

SUMMARY

This report documents the initial field testing and refinement of the officer Leadership Effectiveness Assessment Profile (LEAP), a biographical selection and classification measure being developed as a possible adjunct to the Air Force Officer Qualifying Test (AFOQT). The instrument was revised five times; each revision used a test administration methodology appropriate to its level of development. Earlier iterations used more personalized modes of administration and respondent feedback. Because of the population-specific nature of biodata measures, independent versions of the LEAP were developed for Reserve Officers' Training Corps (ROTC) and Officer Training School (OTS) populations. The latter has received less intensive development and testing than the former, so this report focuses primarily on the ROTC-related measure.

While further refinement, validation and replication of the ROTC instrument are required, considerable progress has been made. The overall test-retest reliability is .73, with scale reliabilities ranging from .48 to .81.

An empirical key was developed for the ROTC instrument to optimize the validity of its scores. Using transformed scores based on the ALS Ordinal empirical key, it was established that when the LEAP was added to the AFOQT, the R^2 increased from .04 to .30 against a composite field training performance criterion.

When the investigators sought to validate the LEAP against a newly developed, 19-dimension peer rating scale, the scales provided only modest support for the validity of the measure.

Analyses were also conducted to ascertain if systematic response bias existed based on gender, ethnicity and socioeconomic status. Subgroup analyses were performed to compare mean LEAP O-2D component scale scores for males and females, whites and non-whites, and high versus low family income respondents. Overall, these subgroup analyses yielded minimal scale score differences, supporting the conclusion of absence of bias.

A second type of response bias was investigated: bias due to social desirability. To determine to what degree, if any, that was occurring, a 12-item Faking Detection scale was developed, piloted and embedded in LEAP 0-2D (ROTC). Results revealed that faking occurred only to a limited degree, and that the faking was confined to only a few of the 14 LEAP scales. The Team Player Orientation scale was particularly vulnerable. A more definitive test of the faking proneness of the LEAP must await a "pre-entry" administration of the LEAP to ROTC or OTS applicants.

Recommendations for further development and field testing of the instrument are presented.

INTRODUCTION

The Armed Services Vocational Aptitude Battery (ASVAB) and the Air Force Officer Qualifying Test (AFOQT) are the primary psychometric vehicles for Air Force personnel selection and classification. They effectively measure general and some specific cognitive abilities. The ASVAB and AFOQT do not, however, measure specific cognitive abilities very well (Morales, 1991; Ree & Earles, 1990a, 1990b, 1990c; Welsh, Watson, & Ree, 1990), nor do they measure personality attributes, psychomotor abilities, leadership or managership potential, biographical information, or how people process information. Thus, the Air Force is investigating new measures using paper-and-pencil and other modes of assessment to enhance personnel selection, classification and related matters (Kyllonen, in press; Berger, Gupta, Berger, & Skinner, 1990b; Carretta, 1987; Driskell & Olmstead, 1989; Siem, 1990; Watson & Besetsny, 1991, 1992; Watson, 1989; Watson, Elliott, & Appel, 1988).

Although Air Force investigators have been developing measures tapping a variety of attributes, the present researchers were concerned with developing measures that would assess leadership potential, managership potential, a propensity for commitment to the Air Force, and related attributes. Two approaches to measuring such attributes were considered: assessment center technology and biodata. Elliott and Watson (1987) evaluated the usefulness of assessment center measures of leadership and managership potential and concluded they had considerable potential. However, these techniques were expensive and labor intensive, which discouraged their use with large numbers of annual applicants to the Air Force. Robertson and Smith (1989), using meta-analytic techniques, synthesized the large amount of data available on the validity of commonly used predictors. Table 1 shows how each is related.

As can be seen in Table 1, biodata appear to be moderately effective as predictors. Other authorities have long supported the use of biodata instruments as a cost-effective methodology for selection purposes (Mumford & Owens, 1987; Owens, 1976; Sparks, 1988). For these reasons, the Air Force chose the biodata alternative as the path to pursue.

An earlier technical report (Appel, Grubb, Shermis, Watson, & Cole, 1990), documented the initial development of a conceptually based biographical instrument called the Leadership Effectiveness Assessment Profile (LEAP). This prototype, the first of five versions to date, was developed for use with officer candidates and was therefore designated LEAP 0-1.

In addition, a parallel item pool was developed on the basis of an elaborate organizational taxonomy for use in the construction of a LEAP E-1 instrument for Air Force enlisted personnel. This taxonomy and item pool (Appel, Grubb, Elder, Leamon, Watson, & Earles, 1991) awaits evidence of the utility of the officer LEAP before being developed further. Nevertheless, the experience gained in item development for this related measure proved most helpful in further refinement of the officer LEAP.

Table 1. Range of Mean Validity Coefficients for Commonly Used Predictors of Work or Business Success

Predictor	Range of Mean Validity Coefficients
Work Sample	.38 to .54
Ability Composite (General Mental Ability plus Psychomotor Ability)	.53
Assessment Center	.41 to .43
Supervisor/Peer Evaluation	.43
General Mental Ability	.25 to .45
Biodata	.24 to .28
References	.17 to .26
Interviews	.14 to .23
Personality Assessment	.15
Self-Evaluation	.15
Interest Assessment	.10

cf. Robertson and Smith (1989)

This report documents the second phase of officer instrument development, in which the LEAP was field tested and revised. Five iterations of testing were accomplished, each focusing on differing psychometric objectives and building on the results of earlier efforts. Each iteration in this refinement process is discussed successively, detailing objectives, methodologies, psychometric properties, and reviews of the LEAP instrument. An overview of the field testing is given in Table 2.

Table 2. Overview of LEAP Field Testing

Version of LEAP	Population Sampled	Sample Size	Location of Sample	Type of Administration	Type of Feedback
O-1	Junior officers	61	Randolph, Brooks AFBs	One-on-one oral	Face-to-face
O-2A	Junior officers	71	Keesler AFB	Small group paper & pencil	Focus groups

Table 2. Concluded

Version of LEAP	Population Sampled	Sample Size	Location of Sample	Type of Administration	Type of Feedback
O-2B (ROTC)	1990 ROTC summer cadets	345	Lackland AFB	Large group paper & pencil	Evaluation questionnaire
O-2B (OTS)	OTS cadets	72	Lackland AFB	Large group paper & pencil	Evaluation questionnaire
O-2C (OTS)	OTS cadets	156	Lackland AFB	Large group paper & pencil	Evaluation questionnaire
O-2D (ROTC)	1991 ROTC summer cadets	673	Lackland, Lowry, McConnell, Plattsburgh, Vandenberg AFBs	Large group paper & pencil	None

METHODS AND RESULTS OF EARLY FIELD TESTING

Field Testing LEAP O-1

The first officer LEAP (LEAP O-1) was developed in a preceding Air Force project. That initial project provided a conceptual model of Air Force officer effectiveness and retention, which served as the basis for generating LEAP item content, and is described in detail elsewhere (Appel et al., 1990). Each of the 12 scales used in LEAP O-1 is briefly defined as follows:

1. Transformational Leadership (Trf Ldr): an approach used by leaders to raise the consciousness of others regarding issues of consequence by effectively arguing for them and thereby mobilizing participation for the good of the team, organization, or polity at levels far beyond what might have been expected.

2. Transactional Leadership (Trn Ldr): a traditional leadership approach characterized by the leader's effort to motivate others by exchanging contingent

rewards or punishments commensurate with the quality and complexity of services rendered. The leader strives to find and provide rewards of the sort desired and thereby enhance performance.

3. Decision-Making Abilities (D-M Abl): those information management skills which permit a leader to effectively evaluate and use job-related information to arrive at decisions.

4. Giving/Seeking Information (G/S Inf): the ability of a leader to give and obtain information necessary to monitor operations and the external environment, to clarify roles and objectives for tasks needing to be completed, and to provide information as needed to relevant others.

5. Team Player Orientation (T-P Or): an ability to function effectively in joint, collaborative efforts with co-workers when attempting problem resolution, as opposed to independent problem-solving.

6. Self-Sufficiency Orientation (S-S Or): an ability to function effectively in independent problem-solving efforts when attempting problem resolution, as opposed to joint, collaborative efforts with other co-workers.

7. Physical Fitness Factors (Phy Fit): refers to an individual's valuing of life long fitness, manifested in a desire for exercise, proper diet, and maintaining good health.

8. Institutional Commitment (Inst Com): refers to a set of attitudes and behaviors of an individual which transcends self-interest to contribute to the success of an organization's mission.

9. Occupational Commitment (Occ Com): refers to a set of attitudes and behaviors which place a higher priority upon the gratification of self-interest than on the interests of the organization with which the individual is affiliated.

10. Persistence to Excellence (Prs Excl): the inclination not to be satisfied with one's level of proficiency until the highest standards of excellence are achieved.

11. Toleration of Adversity (Tol Adv): the ability to endure hardship and frustration without allowing those matters to discourage the individual from the pursuit of his or her goal.

12. Retention Propensity (Ret Prp): the quality and quantity of other employment opportunities which the individual believes are realistically available compared with the position currently held. In early versions (LEAP O-1 through O-2C) quality and quantity of options were evaluated separately, but thereafter were combined into a single component scale.

The composition by scale of the 102-item LEAP O-1 is detailed in Table 3.

Table 3. Composition of LEAP O-1 by Scale

Scales	Number of Items
Transformational Leadership	9
Transactional Leadership	4
Decision-Making Abilities	8
Giving/Seeking Information	7
Team Player Orientation	7
Self-Sufficiency Orientation	4
Physical Fitness Factors	7
Institutional Commitment	17
Occupational Commitment	13
Persistence to Excellence	3
Toleration of Adversity	3
Retention Propensity	5
Classification[a]	9
Demographics[b]	6
TOTAL	102

[a]Classification is a general heading for all questions about college major, academic standing, grades, etc.
[b]Demographics is a general heading for all questions about age, gender, ethnicity/race, socioeconomic status, region, etc. These items were included for research purposes only and were not intended for use in a future operational version of the LEAP.

Objectives

The objectives of the initial LEAP O-1 field testing were to ensure that each item:

1) clearly communicated the intended meaning,
2) was written at a level which respondents could understand,
3) referred to content which respondents could recall, and
4) allowed for complete and thorough answers.

In addition, response alternatives were evaluated to ensure that:

1) they were approximately equal in social desirability,
2) the entire range of potential responses was covered, and,
3) the alternatives were not redundant.

Further, the frequency distribution of response alternatives was examined to ensure adequate variance; items requiring written responses were examined for systematic response patterns from which to generate more objective response alternatives; and item content was screened for insensitive phrasing or terminology objectionable to minority or female respondents.

Subjects

Although the target populations for the LEAP were OTS and ROTC officer applicants, the difficulty of obtaining their participation necessitated the use of an alternative respondent pool. Junior officers on active duty provided an appropriate and accessible respondent population for conducting a pilot test of LEAP O-1.

Participants were 61, second and first Lieutenants and Captains stationed primarily at Randolph AFB. This sample was all of the on-base lieutenants and junior captains that could be assembled for participation. Because subjects were difficult to obtain, three additional subjects were recruited from Brooks AFB. The sample was approximately 75% male and 25% female. Ethnic composition of respondents was as follows: 81% were White; 9.9% were Black; 4.2% were Asian; 2.8% were Hispanic; and 1.4% were American Indian. Reports of parents' total income during respondents' high school years indicated 60.2% in the middle income range ($20,000-$50,000). Finally, approximately 75% of the respondents had spent 3 years or more in the service.

Procedures

A one-on-one oral administration was used to obtain extensive feedback about officers' responses and difficulties encountered on individual items. In LEAP O-1 (and other early versions), individual or group feedback was an important component of instrument refinement since respondents were considered contributory developers of the LEAP.

Each oral administration lasted 2 hours and was conducted in private, air-conditioned, base classrooms. Six specially trained researchers simultaneously administered the instrument over a 3-day period in January 1990. Each biographical item was presented to the respondent on a 5x7 index card and participants' responses were recorded in a separate answer booklet. Included in the instrument were several special follow-up questions designed to address issues of clarity, ease of response, social desirability, and completeness. Table 4 illustrates sample questions for various items with reference to a specific objective. Each follow-up question was also presented on a 5x7 card immediately after the participant responded to the relevant item. For certain questions, respondents were also shown a Likert scale and asked to indicate the level of that particular concern on a 5-point scale. Verbal descriptors, such as "very hard" or "very easy," were used to anchor each end of the scale. The LEAP

administrators also recorded respondents' spontaneous comments and made observations about testing conditions.

Table 4. Sample Follow-Up Questions for LEAP O-1

Objective	Question
Vocabulary:	What does the phrase "agenda" mean to you?
Frame of Reference:	When I said, "redesign your job," what kind of changes came to your mind?
Ability to Recall:	How hard or easy was it for you to remember this information? (show Likert scale)
Clarity:	Was there a single clear answer or did you use several strategies depending on the circumstances?
Comprehen-siveness:	Can you think of another answer (not given here) that would help you answer this question as accurately as possible?
Mutually Exclusive Categories:	Would it be possible for you to answer this question by checking more than one of these categories?
Social Desirability:	Do you think people would be tempted to mark any of these answers over the others? (If so, show Likert scale: How tempted would they be?)
Degree of Threat:	How comfortable or uncomfortable were you about answering this question? (show Likert scale)
Relevance:	To what degree does this question seem related to aspects of the Air Force? (show Likert scale)
Bias of Question:	Was the level of competition at your school affected by its size or quality? Explain.

Analyses

For open-ended questions, individual responses were used to identify appropriate response alternatives and objective items were developed. Similarly, the 23 behavioral grid items were each analyzed and translated into an objective format. A distractor analysis--an analysis of the frequency distribution of response alternatives--was performed for each objective item. The mean and range of

the follow-up questions were also computed to check for any indications of confusion, social desirability, etc.

Results

Approximately 70 items were edited or substantially changed after the data and respondent feedback were analyzed. Within each item, response alternatives were modified, discarded, or generated to encourage a wide range of response. The following examples illustrate typical modifications. For the item "Check the college grade you most often received in the following subjects," several subject areas were added to make the response alternatives inclusive. Low variability (87% yes, 13% no) led to removal of the item "Did you attend your high school graduation?" Table 5 shows an example of the evolution of an item over successive LEAP administrations. The item shown was modified in early versions but remained the same in later versions as reasonable and stable percentages were obtained for each response option across multiple administrations.

A second LEAP, designated LEAP O-2A since it was the first version developed under the second LEAP contract, was subsequently constructed and forwarded to the Air Force Human Resources Laboratory (AFHRL[1]) for review. The suggestions of AFHRL scientists were incorporated, resulting in a 91-item, objectively formatted instrument.

Field Testing LEAP O-2A

Objectives

LEAP O-1 field testing helped determine more specific objectives for LEAP O-2A. Now all in objective format, the items were examined for appropriate response alternative frequencies. Respondent feedback was still a vital part of instrument refinement; so, an evaluative questionnaire and small focus groups were used to test respondent ability to recall past events, lack of specificity in response alternatives, clarity of particular words and phrases, and any other ambiguities or difficulties. A third objective was to determine if respondents were "gaming" the instrument--that is, selecting a socially desirable response alternative. A final objective of LEAP O-2A field testing was to determine if the measure could be completed in less than 1 hour.

The composition of LEAP O-2A appears in Table 6. Though there is little change from the LEAP O-1 instrument in the number of items per scale, considerable editing of the items resulted from the review process.

[1] AFHRL has been redesignated the Human Resources Directorate, Armstrong Laboratory.

Table 5. Successive Versions of LEAP Item

Version of LEAP	Percentage of Response	Item Content
O-1		In the past, my typical response to stress in group situations has been to:
	43.7%	A. inject humor into the situation.
	7.0%	B. try to ignore the tension.
	1.4%	C. put the task aside and do something else.
	14.1%	D. find some way to relax myself.
	33.8%	E. openly discuss the tension.
O-2A		In stressful situations within a group, typically my first response has been to:
	45.9%	A. inject humor into the situation.
	4.9%	B. try to ignore the tension.
	3.3%	C. put the task aside and do something else.
	13.1%	D. find some way to relax myself.
	21.5%	E. openly discuss the tension.
O-2B		My first response to stressful group situations has typically been to:
	19.1%	A. find some way to relax myself.
	23.5%	B. try to work despite the tension.
	15.7%	C. openly discuss the tension.
	41.7%	D. inject humor into the situation.
O-2C		My first response to stressful group situations has typically been to:
	24.5%	A. find some way to relax myself.
	23.0%	B. try to work despite the tension.
	12.9%	C. openly discuss the tension.
	39.6%	D. inject humor into the situation.
O-2D		My first response to stressful group situations has typically been to:
	24.3%	A. find some way to relax myself.
	20.6%	B. try to work despite the tension.
	18.3%	C. openly discuss the tension.
	36.9%	D. inject humor into the situation.

Table 6. Composition of LEAP O-2A by Scale

Scales	Number of Items
Transformational Leadership	6
Transactional Leadership	3
Decision-Making Abilities	4
Giving/Seeking Information	6
Team Player Orientation	5
Self-Sufficiency Orientation	3
Physical Fitness Factors	7
Institutional Commitment	14
Occupational Commitment	9
Persistence to Excellence	2
Toleration of Adversity	2
Retention Propensity	4
Classification	15
Demographics	11
TOTAL	91

Subjects

Respondents were 71 junior officers at Keesler AFB, Biloxi, Mississippi. A training center for an array of Air Force occupational specialties, Keesler AFB, was selected because it afforded a large supply of second and first lieutenants.

Also, both rated and non-rated officers were represented. The respondents were primarily male (77%), White (74%), first and second lieutenants participating in initial training within their Air Force specialty. Approximately 20% of this group were health professionals who entered the Air Force through direct commissions.

Procedures

LEAP O-2A was administered in mid-March, 1990 to seven groups of eight to twelve respondents. In addition to LEAP O-2A, participants filled out an evaluation questionnaire asking them to identify ambiguous terms and to evaluate the clarity and relevance of O-2A questions. Respondents suggested additional background information which might be used to identify qualified officer applicants. Respondents also met in small focus groups for half an hour to offer additional reactions to LEAP O-2A.

Analyses

A distractor analysis was generated by calculating the frequency distribution of the item responses. Also, to test the hypothesis that more experienced junior officers might be "gaming" the instrument, a contingency table analysis was conducted comparing responses of naive and experienced officers.

Results

Based on the distractor analysis, response alternatives with low response frequencies were changed, eliminated, or combined with other alternatives. In the contingency table analysis, few significant differences in the response distributions of naive versus experienced officers were found, indicating that knowledge of military customs and procedures did not significantly affect responses.

A new version of the instrument, developed on the basis of these analyses and the focus group feedback, was sent to AL/HR for review; necessary changes were made. This 3-month effort resulted in LEAP O-2B, an 84-item objective instrument now divided into three parts: Part I (demographic information), Part II (instrument), and Part III (brief evaluative questionnaire).

Field Testing LEAP O-2B

Objectives

The next field testing cycle on LEAP O-2B involved both a small sample of Officer Training School (OTS) cadets, and a large sample of Reserve Officers' Training Corps (ROTC) cadets. Differences in age and experience (e.g., OTS college graduates ranging in age from 22 to 30 versus 19- to 20-year-old ROTC college students) led to the development of a special ROTC version in which all questions were worded to reflect cadets' circumstances.

The objective for the smaller OTS sample was to establish the reliability of the LEAP instrument. Achieving a strong reliability coefficient for the measure was particularly important since reliability places an upper limit on validity. For example, if the reliability coefficient is $r = .64$, then the highest possible validity coefficient would be $r = .80$.

However, estimation of the reliability of biodata survey forms can be particularly troublesome. As Mumford and Owens (1987) have pointed out, biodata are more appropriately evaluated for reliability by a coefficient of stability than by an internal consistency measure such as Cronbach's alpha:

> The relative independence of background data has certain implications for the assessment of reliability. More specifically, the independence of these items makes it unlikely that the resulting scales will yield high internal consistency coefficients. Therefore, it is not surprising that the

internal consistency coefficients obtained for rational background data scales lie between .40 and .80. Yet as the verticality studies would imply, background data items commonly yield substantial retest reliability coefficients. For instance, Bunch (1974) obtained retest reliabilities of .60 and .80. Similarly, Saunders (1983) obtained retest coefficients near .60 in correlating item responses at age 18 and 22. Thus, it appears that background data items provide an unusually reliable description of differential behavior and experiences, even over relatively long intervals (p. 7).

The objective for the large sample was to carry out an initial validation of the measure. LEAP O-2B total scores for respondents would be correlated against an overall training performance rating made for each ROTC cadet by supervisory staff at the end of a 4-week summer encampment.

The composition of LEAP O-2B is outlined in Table 7. The minimum number of items per scale was increased to seven (except Retention Propensity, which had six), to enhance reliability in further studies. Classification and Demographic items were moved to Part I and were administered prior to the scale items.

Table 7. Composition of LEAP O-2B by Scale

Scales	Number of Items
Transformational Leadership	7
Transactional Leadership	7
Decision-Making Abilities	7
Giving/Seeking Information	7
Team Player Orientation	7
Self-Sufficiency Orientation	7
Physical Fitness Factors	7
Institutional Commitment	7
Occupational Commitment	7
Persistence to Excellence	7
Toleration of Adversity	8
Retention Propensity	6
TOTAL	84

Subjects

The test-retest objective necessitated the availability of an appropriate sample of junior officers who would be available for two administrations of LEAP O-2B at least 1 month apart. Respondents were 72 cadets from OTS class 91-03, participating in a 12-week training program at Lackland AFB. For the initial validation objective, LEAP 0-2B was administered to 344 ROTC cadets attending one of three 4-week 1990 summer encampments at Lackland AFB, Texas.

These cadets were drawn from a variety of university ROTC programs across the country. Most had just completed their sophomore (second) collegiate year. One hundred fourteen ROTC cadets were assessed in the first encampment, an additional 132 in the second, and 98 in the third. ROTC cadet respondents were primarily male, white, 19- to 20-year-old, middle-class college students from the Southwest.

Procedures

LEAP O-2B (OTS) was group administered to OTS class 91-03 on 19 October 1990, and retested on 7 December 1990. On both occasions, all respondents completed the instrument.

LEAP O-2B (ROTC) was group administered on three occasions (June, July, August 1990) to ROTC cadets. Respondent feedback was noted on a questionnaire at the end of the instrument. Data from the three groups were combined into a single ROTC database. The number of cadets in the sample (n = 344) was sufficiently large to permit a stable assessment of the LEAP's psychometric properties.

Analyses

A test-retest analysis of the LEAP O-2B was conducted for the OTS sample. Also, for the first time, criterion data were available to permit preliminary validation of LEAP O-2B with the larger sample of ROTC cadets. Scores, using the rational key,[2] were computed for each of the component scales and for the total LEAP. These scores were subsequently correlated with a composite ROTC training performance criterion score (see Appendix C).

The criterion was an overall field training performance rating which placed each cadet into one of four quartile groups. The rating was made by ROTC summer encampment faculty and staff. They combined evaluations on ten facets of cadet training performance into a single global score for each cadet. The ten performance factors (see pages 44-45 for definitions) included:

Adaptability to Military Training
Duty Performance
Leadership/Followership
Adaptability to Stress

[2] The rational key was developed on the basis of subjective assessments of the merits of each response alternative to each LEAP item. Merit was judged by the degree to which each response alternative implemented the conceptual model. The most appropriate response alternative in each set was assigned a value of 1.00. Remaining response options were assigned lesser, fractional weights. Other options seen as having no merit were assigned a value of 0.

Drill and Ceremonies
Human Relations
Physical Fitness
Communication Skills
Judgment and Decisions
Professional Qualities

Unfortunately, only composite ratings and not component scores were available for the LEAP O2-B sample. However, component performance factor ratings were obtained for the later, LEAP O-2D sample, and used in analysis at that level of instrument development.

Cadets were ranked by their composite score into first, second, third and fourth quartile groupings. The quartile ratings were correlated with their corresponding total LEAP scores. This provided initial evidence about LEAP O-2B's predictive efficiency.

A correlational analysis was used to identify the degree to which respondent endorsements of particular response alternatives were associated with each quartile. The investigators reasoned that if the LEAP construct scales were operating as desired, there should be a linear relationship between quartile group and each LEAP scale score; that is, cadets in the top quartile would be more likely to select the preferred rational keyed response alternatives than would cadets in the bottom quartile.

Results

The test-retest results obtained for the LEAP O-2B OTS sample appear in Table 15 (page 28) along with test-retest results from later LEAP versions. Note that the test-retest reliabilities varied widely, from a low of .15 to a high of .81, with a test-retest reliability of .64 for the total LEAP score. Test-retest estimates of reliability for three of the component scales, Transactional Leadership (.15), Institutional Commitment (.31), and Occupational Commitment (.31), were unacceptably low. Four other coefficients were marginally low, falling in the .40 to .50 range. Clearly, revision of these seven scales was required.

Preliminary validation results from the correlational analysis are presented in Table 8. Note that there was a low, but statistically significant, relationship between 8 of the 12 component LEAP scales and the ROTC composite training performance criterion. Self-Sufficiency Orientation ($r = .23$), Institutional Commitment ($r = .16$), Physical Fitness Factors ($r = .15$), and Giving/Seeking Information ($r = .15$) are the most valid scales. However, despite the modest correlations with this rather weak criterion, the results indicated that further scale refinement was necessary.

Table 8. Degree of Association Between LEAP O-2B Scale Scores and Quartile Groupings for ROTC Cadets[a]

LEAP Scales	Correlation with Quartile Group
TOTAL LEAP	.25***
Transformational Leadership	.10**[b]
Transactional Leadership	.03
Decision-Making Abilities	.09
Giving/Seeking Information	.15**
Team Player Orientation	.12*
Self-Sufficiency Orientation	.23***
Physical Fitness Factors	.15**
Institutional Commitment	.16**
Occupational Commitment	-.08
Persistence to Excellence	.13*
Toleration of Adversity	.07
Retention Propensity	
Quantity of Alternatives	-.05
Quality of Alternatives	.10*

[a] n = 331
[b] where p is: * = <.05, ** = <.01, *** = <.001.

Field Testing LEAP O-2C

LEAP O-2C, the third LEAP revision, was constructed using item, scale and other psychometric data from the previous field tests. It also incorporated appreciable editorial suggestions on item wording and content offered by AL/HR scientists and external biodata consultants, Drs. C. Paul Sparks and William A. Owens, who provided extensive feedback. LEAP O-2C retained the three-part format of LEAP O-2B (a demographic characteristics section, the instrument, a brief evaluative questionnaire) but, with 86 items, was two items longer. The composition of LEAP O-2C appears in Table 9.

Objectives

The first objective of this field testing was to establish the test-retest reliability of LEAP O-2C. A second objective was to validate LEAP O-2C against OTS training performance criteria. The anticipated criteria, however, were not released by OTS personnel; so, this validation objective could not be accomplished.

Subjects

The instrument was administered to cadets from two Lackland AFB Officer Training School (OTS) classes, 91-04 and 91-05. The former was composed of 69 cadets, and the latter, 87. The demographic characteristics of this combined sample are presented in Appendix A, rather than the text, due to the more extensive attributes described. They reflect a somewhat heterogeneous group in two important regards: a wide age range (22 years to over 30 years); and substantial prior enlisted service (47%).

Table 9. Composition of LEAP O-2C by Scale

Scales	Number of Items
Transformational Leadership	7
Transactional Leadership	7
Decision-Making Abilities	7
Giving/Seeking Information	7
Team Player Orientation	7
Self-Sufficiency Orientation	7
Physical Fitness Factors	8
Institutional Commitment	7
Occupational Commitment	7
Persistence to Excellence	7
Toleration of Adversity	8
Retention Propensity	7
TOTAL	86

Procedures

LEAP O-2C was administered to each OTS class in a single, group administration, and cadets completed an open-ended evaluation form. OTS class 91-04 was administered LEAP O-2C on 11 January 1991. A follow-up administration took place on 2 March 1991. On the same date, OTS class 91-05 took its first administration of LEAP O-2C and was retested on 30 May 1991.

Analyses

The measure was scored for component and total scores, and descriptive statistics were computed. A test-retest analysis was conducted using a 2-month interval.

Results

Descriptive statistics for each of the LEAP O-2C scales are presented in Table 10. These results are based upon data gathered from both OTS classes combined into a composite sample of 156 cadets.

Table 10. Descriptive Statistics for LEAP O-2C (OTS)[a]

Scale	Number of items	Mean	Standard deviation	Minimum score obtained	Maximum score possible	Maximum score obtained
TrfLdr[b]	7	3.77	1.22	.66	7.00	7.00
TrnLdr	7	.67	1.09	-2.50	3.00[c]	3.00
D-MAbl	7	3.97	1.08	1.66	7.00	6.33
G/SInf	7	3.99	.91	1.41	7.00	6.41
T-POr	7	3.83	1.16	.50	7.00	6.00
S-SOr	7	4.24	.87	1.83	7.00	5.91
PhyFit	8	5.22	1.26	1.78	8.00	7.75
InstCom	7	4.65	.72	3.27	7.00	6.41
OccCom	7	-2.65	1.07	-5.32	-7.00	0.00
PrsExcl	7	3.30	.81	1.65	7.00	5.08
TolAdv	8	3.96	1.31	1.00	8.00	8.00
RetPrp	7	3.13	1.22	.75	7.00	6.50

[a]n = 141
[b]TrfLdr = Transformational Leadership
TrnLdr = Transactional Leadership
D-MAbl = Decision-Making Abilities
G/SInf = Giving/Seeking Information
T-POr = Team Player Orientation
S-SOr = Self-Sufficiency Orientation
PhyFit = Physical Fitness Factors
InstCom = Institutional Commitment
OccCom = Occupational Commitment
PrsExcl = Persistence to Excellence
TolAdv = Toleration of Adversity
RetPrp = Retention Propensity
[c]Although there are seven items in this scale, the maximum possible score is 3.00 and the minimum possible score is -4.0 because four of the items were assigned negative weights in the rational key.
Note: The minimum possible score for all other scales is 0.

All but four of the component scales (Institutional Commitment, Persistence to Excellence, Giving/Seeking Information, and Self-Sufficiency Orientation) showed good variance, having standard deviations over 1.0. The four scales with more limited variance deal with attributes on which this sample might be expected to score highly, and so produce few scores in the lower end of the scale. In addition, the stability of responses on LEAP O-2C was calculated.

The results are presented in Table 15 jointly with those of LEAP 0-2B and LEAP O-2D measures. Note that only three of the 14 scales (Transactional Leadership, Transformational Leadership, and Occupational Commitment) yielded unacceptably low test-retest reliability coefficients (.20, .46, .47 respectively), and that the overall reliability was improved from .64 to .69. Given that the test-retest interval for LEAP O-2C was twice as long (2 months) as that for LEAP O-2B, this represents an improvement in item quality.

METHODS AND RESULTS OF LATER FIELD TESTING

Field Testing LEAP O-2D (ROTC)

Objectives

Objectives for LEAP O-2D were the most comprehensive of all the versions. Availability of a larger sample (n = 673) and multiple criteria contributed to meeting the following objectives: development of an empirical key, computation of descriptive statistics, further assessment of test-retest reliability, extensive testing of validity hypotheses using three criteria, analysis of intercorrelations among LEAP scales, testing for response bias, and development/testing of a Faking Detection scale.

Also, a number of refinements were made in the construction of LEAP O-2D. The Transactional Leadership scale was modified by eliminating Management by Exception items and adding more Contingent Reward items so all scale items would be positively weighted. The Occupational Commitment scale was deleted since the Institutional Commitment scale alone provided the desired emphasis on moral commitment. Two complementary scales, Quality and Quantity of Work Alternatives, were consolidated into a single scale, Retention Propensity.

Most importantly, two new scales, Charisma and Socialized Power, were added. It has long been noted that Charisma is the most potent element of Transformational Leadership (Bass, 1985). Recent work by Conger and Kanungo (1988) and Conger (1989) have helped to elaborate and operationalize this complex concept. Their work suggests that this construct should stand alone rather than merely function as one of three elements of Transformational Leadership. The addition of the Charisma scale allowed both possibilities to be tested. Similarly, recent research (Winter, 1987) has demonstrated the predictive efficiency of Socialized Power as a component of military leadership. It was included to incorporate a variable not apparent at the outset of the project. LEAP 0-2D emerged as a 137-item measure as indicated in Table 11.

Table 11. Composition of LEAP O-2D (ROTC) by Scale

Scales	Number of Items
Transformational Leadership	22
Charisma[a]	(15)
Transactional Leadership	8
Decision-Making Abilities	7
Giving/Seeking Information	7
Team Player Orientation	7
Self-Sufficiency Orientation	7
Physical Fitness Factors	9
Institutional Commitment	7
Persistence to Excellence	7
Toleration of Adversity	8
Socialized Power	12
Retention Propensity	7
Faking Detection	12
TOTAL	120

[a]Charisma was analyzed both as a part of Transformational Leadership and separately as an independent construct. The 15 items in this scale are also part of the Transformational Leadership scale.

Finally, as with LEAP O-2B, two versions of LEAP O-2D were created: one adapted for ROTC cadets, and one adapted for OTS cadets. However, due to limited subject availability, only the ROTC version was administered.

Subjects

Arrangements were made with Headquarters, Air Force ROTC, Maxwell AFB, to administer LEAP O-2D to approximately 150 ROTC cadets attending each of five summer encampments at McConnell, Lowry, Vandenberg, Lackland, and Plattsburgh AFBs. Approximately 673 cadets served as respondents. However, sample sizes for the various analyses were considerably smaller, clustering about n = 263. A number of respondents were unable to complete the instrument in the allocated testing time; others chose not to respond to particular items; and the answer sheets of some others could not be machine processed.

Demographic characteristics of this composite sample are detailed in Appendix A. In overview, the sample was predominantly a homogeneous group of 19- to 21-year-old white, single, male, ROTC college students from urban/suburban settings.

Procedures

LEAP O-2D (ROTC) was initially administered during July 1991 by resident faculty at each of the encampments. Instructions, copies of the instrument, answer sheets, and other necessary materials were sent to ROTC training officers, who administered the measure and returned completed results to the Armstrong Laboratory. To enable test-retest reliability analysis of the new instrument, ROTC training officers readministered LEAP O-2D during August 1991, 3 to 4 weeks after the initial administration.

Following completion of the retest administration, cadets at two of the five encampments, Lackland AFB and Plattsburgh AFB, participated in a peer rating analysis. Cadet flight groups were divided into two groups of 10 each, and each cadet rated 8 of the 10 cadets in the alternate group. The assumption was made that the raters knew their flightmates well enough to provide dependable evaluations. A peer rating instrument was administered using procedures developed by Armstrong Laboratory and LEAP personnel. The AFROTC Peer Rating Form (AFPRF) was constructed specifically for the LEAP project and contained 19 dimensions (see Appendix E). Seventeen of these dimensions were designed to provide peer rating criteria specific to the constructs of the LEAP. In most instances, each dimension taps a single construct. However, for three of the more complex scales, more than one dimension was required. In these three instances (Transformational Leadership, Decision-Making Abilities, and Socialized Power) a composite ratings score was generated. In addition, two of the dimensions, #18 and #19, provide more global performance criteria: encampment success and future potential, respectively. Each dimension was presented using a five-point rating scale which also included a not-enough-information-to-respond category. Respondents rated their peers on these 19 dimensions, generating an independent set of criteria against which to validate LEAP O-2D.

Analyses and Results

All analyses performed on LEAP O-2D (except Faking Detection scale analyses) were based primarily upon empirical, rather than rational, keyed data. The empirical key was based on an alternating least squares (ALS) Ordinal algorithm developed by Young (1981). Because of the complex procedures entailed in its development, and the importance of the empirical key as a basis for most analyses, the next sections introduce the methodology used in its construction, and the rationale for employing the ALS Ordinal approach over alternate available empirical key-building strategies. Results of the analyses of descriptive statistics, possible biases, reliability, and validity appear in individual sections following the sections on empirical key development.

Development of an Empirical Key

The scoring strategy used in the LEAP assumes that some of the response alternatives (besides the one best answer) have varying degrees of merit. Hence, for several items, partial credit is given for partially correct response alternatives. For example, a question with four response alternatives a, b, c and d might be scored as 1.0, .75, .50, and 0, respectively. This contrasts with the common multiple-choice scoring schema in which a single response alternative is scored as 1.0 and all others as 0.

It is possible to generate a logic for specifying correct and partially correct responses, as the investigators did in developing the rational key. However, it is generally not possible to decide, *a priori*, the precise weights for partially correct response alternatives (e.g., .75 or .66). Fortunately, methodologies exist which permit an instrument designer to optimize item or scale scores by empirically deriving weights for item response alternatives. Use of such optimization procedures can sometimes markedly improve the efficiency of a predictive instrument over that possible using a rational, or subjectively derived key.

Existing item/scale score optimization procedures were examined, narrowing the available options to three. These three, ALS Nominal (Young, 1981), ALS Ordinal (Young, 1981), and Correspondence Analysis (Greenacre, 1984), were then used in generating parallel empirical keys, and applying those keys to randomly generated (Monte Carlo) data and also to data from the LEAP O-2D administration. This parallel effort was used to determine which of the three approaches was most effective. Appendix B describes the procedures required to generate empirical keys by each of these three methods.

Both the Nominal and Ordinal ALS algorithms work by dividing all of the item distribution statistics into two mutually exclusive subsets: (a) the parameters of the model; and (b) the parameters of the data (i.e., the optimal scaling parameters). The algorithms then optimize a loss function by alternately optimizing with respect to one subset, then the other. The optimization proceeds by obtaining the least squares estimates of the parameters in one subset while assuming that the parameters in all other subsets are constants. Once a conditional least squares estimate has been obtained, the old estimates of the parameters are replaced by the new estimates. The algorithm then switches to another subset of parameters (i.e., each of the two subsets may itself contain parameters subsets) to obtain their conditional least squares estimates. The iterations continue until convergence takes place.

The main difference between the ALS Ordinal and Nominal algorithms is that with the Ordinal approach a "correct" answer can be designated and the remaining responses empirically reweighted. This approach is useful if it is desirable to maintain a correspondence with the rational key. In the case of the ALS Nominal, there are no constraints. All of the response alternatives are free to be empirically reweighted without reference to the rational key.

Correspondence Analysis provides optimal weights based on the dimensionality of a predictor and its criterion. This approach is grounded on the fundamental singular values decomposition of a matrix and has been alternatively referred to as optimal scaling, dual scaling, Guttman Scaling, and Pattern Analysis (Weller & Romney, 1990). The first step in Correspondence Analysis is to normalize the data by dividing each row entry by the square root of the product of corresponding row and column totals. In the second step, the basic structure of the normalized matrix is found using the singular value decomposition (SVD) technique. The last step is to rescale the row and column vectors to obtain the canonical or optimal scores. Correspondence Analysis essentially treats all the data as if they were Nominal.

A decision was made to use the ALS Ordinal empirical key in preference to the other two. Using this approach, the "correct" answer could be based on the rational key, but the weights given to the "secondarily correct" and to "incorrect" answers could be determined empirically. This approach was the only one which allowed the desired compromise: maintaining the "seed" response as dictated by the LEAP's conceptual framework, and allowing empirical results to estimate weights for all other response alternatives.

As an example of the relationship among the several item keying approaches explored, here are representative empirical key results achieved for one item on the LEAP instrument: "What kind of appointment book or calendar do you keep?" The weights computed or assigned for each of the response alternatives are given in Table 12. Note that the Nominal key mean criterion score for respondents endorsing alternative "B" was higher than that for respondents selecting alternative "A" (3.0 vs. 2.86). This pattern of response is not consistent with the pattern of weights as dictated by the Rational key, where alternative "A" had been assigned the higher weight. This conflicting outcome underscores the fact that the nominal keying approach is based purely on empirical weights, unbounded by theoretical constraints.

In the Ordinal key, however, keyed results bounded the empirical weights with an order dictated by the theory. This required that the weight for alternative "A" must be at least as high, if not higher, than that for alternative "B." Similarly, the weight derived for alternative "B" must be as high, if not higher, than that computed for "C." Correspondingly, the weight for alternative "C" must be as high, if not higher than the weight for alternative "D." Note that the Ordinal key weights incorporate nominal values when those values are consistent with the order dictated by the rational key.

Thus, it can be seen that the Ordinal key functions as a hybrid. Consequently, alternative "A" and "B" both received a compromise weight of 2.95. The weight for alternative "A" is less than that called for by the nominal key, but meeting the minimal requirement of the rational key; and the empirically derived weight of 3.0 for alternative "B" is reduced to 2.95 to meet the constraint of the theoretical model.

Table 12. Comparative Weights for Response Alternatives on an Illustrative O-2D Item Using Rational, Ordinal, and Nominal Keying

Item: What kind of appointment book or calendar do you keep?			
Item Response Alternative	**Rational Key**	**Ordinal Key**	**Nominal Key**
A. A meticulous record of present and future events	1.00	2.95	2.86
B. A record of important future events	.66	2.95	3.00
C. A simple calendar of future events	.33	2.82	2.82
D. I have never kept one	0.00	2.73	2.73

Examining both empirical and rational keys has the advantage of providing a basis for reconsidering the original theoretical rationale. For example, the Nominal key results invite a re-examination of the assumption that "A meticulous record of present and future events" is optimal. Perhaps that is excessive record keeping. Maintaining "A record of important future events" may be a more economical and sufficient strategy for effective officer performance.

Descriptive Statistics Results

Descriptive statistics for LEAP O-2D are presented in Tables 13 and 14. The data are presented in parallel form: Table 13 contains rational key results; Table 14 contains corresponding Ordinal key outcomes. Results were combined from all five ROTC summer encampments into a single composite sample. A comparison of means and standard deviation values for the component scales on both sets of data must take into consideration that the length of the component scales ranges from a minimum of seven to as many as 23 items. Hence, only scales with the same number of items are directly comparable.

The sample for the Ordinal empirical key (n = 263) is based on a much smaller sample than for the rational key data (n = 518-612). This is attributable to the fact that the Ordinal empirical key transformation required complete data for a participant to be included in the data set.

Some clarification is required in explaining the minuscule standard deviations produced by the Ordinal key data as given in Table 14. Those results may

be attributed to at least four factors, each of which contributed to limiting variance. The first consideration is the large number of items set to a constant value as a result of a mismatch between rational and Nominal keys. In this mismatch, the preferred response alternatives as dictated by the rational key are in direct conflict with the response alternatives actually selected by respondents. The second factor is that for a substantial number of items not all of the response alternatives functioned as distractors. The third factor was the similarity of actual weights assigned by the Ordinal optimizing procedure to each response alternative, restricting variance. Fourth, restriction in criterion range was due to limited use of extreme values in the four-point rating scale (see Table 16, p. 30).

Test-Retest Reliability Results

For ease of comparison, results of the test-retest analyses conducted on LEAP O-2B, O-2C (OTS), and O-2D (ROTC) are presented together in Table 15. Note that the time interval of the three administrations differed, ranging from 1 to 2 months. Also, for LEAP O-2D, results are presented using both rational and Ordinal key data. Overall, the reliability coefficients show general improvement across the three versions of the LEAP instrument: .64 for LEAP O-2B, .69 for LEAP O-2C, and .73 for LEAP O-2D, respectively. Variability among the component scale reliabilities is detailed next.

Whereas LEAP O-2B produced coefficients for some scales as low as .15, with the highest reliability no larger than .81, LEAP O-2D yielded reliability coefficients ranging from .48 to .81. In part, this enhanced reliability is attributable to the increased length of the measure, which grew from 84 items to 137, and also to scale refinements. It is particularly noteworthy that the three new LEAP scales (Charisma, Socialized Power, and Faking Detection) achieved an acceptable level of reliability in their first field testing.

As may be seen in Table 15, a fourth test-retest analysis was carried out using the Ordinal empirical key. Note that the reliability coefficients are generally lower than those achieved for the rational key. It is hypothesized that this outcome is attributable to the nature of the optimizing procedure, which maximizes the relationship with the criterion measure at the possible expense of test-retest reliability.

Validating LEAP O-2D

The approach to validation taken in this study is based on the work of Messick (1989). Messick argues that the traditional division into content, construct, concurrent, and predictive validity is outdated. He postulates that validation of an assessment instrument involves hypothesizing the existence of certain relationships between a construct of interest and the criterion, followed by the collection of data to test the hypotheses.

Table 13. Descriptive Statistics for LEAP O2D, Based on Rational Key Data[a]

Scale	N of Items	Mean	Standard Deviation	Minimum Score Obtained	Maximum Score Possible	Maximum Score Obtained
Tot LEAP[b]	120	5.91	1.06	3.48	10.00	6.66
TrfLdr[c]	23	9.79	2.36	2.99	23.00	17.15
Chrs	16	6.22	1.57	2.33	16.00	10.99
TrnLdr	8	3.96	1.08	0.50	8.00	6.91
D-MAbl	7	4.57	0.98	1.33	7.00	7.00
G/SInf	7	4.24	0.86	1.81	7.00	7.00
T-POr	7	3.75	1.13	0.00	7.00	6.66
S-SOr	7	4.16	0.97	1.00	7.00	6.66
PhyFit	9	5.54	1.20	0.75	9.00	8.55
InstCom	7	4.31	0.88	1.32	7.00	6.66
PrsExcl	7	4.25	0.99	0.70	7.00	7.00
TolAdv	8	4.55	1.24	1.41	8.00	8.00
SocPwr	12	5.64	1.45	2.00	12.00	11.00
RetPrp	7	3.27	1.69	0.25	7.00	5.75
FakDet	12	4.25	1.69	0.00	12.00	9.00

[a]Sample size varies due to missing data. The range is from 518 to 612.
[b]TotLEAP = Total LEAP score; a composite across 12 weighted component scale scores (not including Charisma and Faking Detection).
[c]TrfLdr = Transformational Leadership
Chrs = Charisma
TrnLdr = Transactional Leadership
D-MAbl = Decision-Making Abilities
G/SInf = Giving/Seeking Information
T-POr = Team Player Orientation
S-SOr = Self-Sufficiency Orientation
PhyFit = Physical Fitness Factors
InstCom = Institutional Commitment
PrsExcl = Persistence to Excellence
TolAdv = Toleration of Adversity
SocPwr = Socialized Power
RetPrp = Retention Propensity
FakDet = Faking Detection
Note: The minimum possible score for all scales was 0.

Messick goes on to say that the validity of a measure is related to the particular uses of that instrument. Thus, validity is not a general attribute but a function that varies according to each of the purposes for which the instrument is designed. As a result, an instrument's validities will vary with its applications. From the above, the need for defining intended applications of any measure

can be seen. This is particularly true in the case of a newly established instrument such as the LEAP. As a first application, the designers sought to use the LEAP in conjunction with the AFOQT, to enhance predictive efficiency in the selection of ROTC and OTS cadets. With subsequent development of its component scales, the LEAP may have a second application as a diagnostic tool, focusing upon a cadet's needed areas of training or identifying a cadet's strengths for classification purposes. Finally, there was concern about the applicability of the LEAP to non-traditional Air Force applicant groups. It was important to establish that this new measure was as applicable to females, non-whites, and low socioeconomic status respondents as it was to the traditional white, male, economically advantaged applicant. In a closely related concern, the researchers needed to establish the absence of any substantial response bias by respondents seeking to distort their answers in a favorable direction.

Table 14. Descriptive Statistics for LEAP O-2D, Based on Ordinal Key Data[a]

Scale	N of Items	Mean	Standard Deviation	Minimum Score Possible	Minimum Score Obtained	Maximum Score Possible	Maximum Score Obtained
Tot-LEAP[b]	120	302.22	.80	297.58	299.82	307.21	304.24
TrfLdr	23	57.57	.27	55.20	56.91	59.31	58.79
Chrs	16	37.42	.18	36.94	37.04	38.33	8.02
TrnLdr	8	20.15	.08	20.07	20.07	20.45	20.41
D-MAbl	7	20.15	.15	19.42	19.32	20.32	20.29
G/SInf	7	20.15	.14	19.16	19.68	20.42	20.41
T-POr	7	20.15	.11	19.94	19.95	20.58	20.39
S-SOr	7	20.15	.15	19.50	19.65	20.49	20.32
PhyFit	9	25.90	.28	24.81	24.95	26.55	26.52
InstCom	7	20.15	.13	19.48	19.77	20.43	20.39
PrsExcl	7	20.15	.18	19.70	19.75	20.43	20.43
TolAdv	8	23.03	.09	22.90	22.94	23.16	23.16
SocPwr	12	34.54	.12	34.17	34.20	34.84	34.73
RetPrp	7	20.15	.15	19.43	19.77	20.82	20.67

[a] n = 263
[b] TotLEAP = Total LEAP Score
TrfLdr = Transformational Leadership
Chrs = Charisma
TrnLdr = Transactional Leadership
D-MAbl = Decision-Making Abilities
G/SInf = Giving/Seeking Information
T-POr = Team Player Orientation
S-SOr = Self-Sufficiency Orientation
PhyFit = Physical Fitness Factors
InstCom = Institutional Commitment
PrsExcl = Persistence to Excellence
TolAdv = Toleration of Adversity
SocPwr = Socialized Power
RetPrp = Retention Propensity

Table 15. Test-Retest Reliability for LEAP O-2B, O-2C, and O-2D

Scale	O-2B (n=72)	O-2C (n=156)	O-2D (n=430)	O-2D[a] (n=263)
Total LEAP Score	.64	.69	.73	.71
Transformational Leadership	.60	.46	.65	.46
Charisma	--	--	.57	.41
Transactional Leadership	.15	.20	.48	.48
Decision-Making Abilities	.57	.55	.63	.67
Giving/Seeking Information	.48	.67	.54	.66
Team Player Orientation	.45	.70	.61	.54
Self-Sufficiency Orientation	.71	.58	.63	.49
Physical Fitness Factors	.49	.80	.71	.63
Institutional Commitment	.31	.59	.67	.66
Occupational Commitment	.31	.47	--	--
Persistence to Excellence	.80	.83	.81	.78
Toleration of Adversity	.47	.65	.63	.64
Socialized Power	--	--	.58	.58
Retention Propensity	--	--	.79	.66
Quantity of Work Alternatives	.81	.84	--	--
Quality of Work Alternatives	.46	.82	--	--
Faking Detection	--	--	.65	.43
Time Interval (in months)	*1*	*2*	*1*	*1*

[a]Unlike the other three data sets, these results are based on the ALS Ordinal (i.e., the empirical) rather than the Rational key.

This section of the technical report is organized into three parts which present evidence for the following: 1) the efficacy of using LEAP O-2D together with the AFOQT as joint predictors of cadet training performance, 2) the utility of the LEAP's component scales for diagnostic or classification purposes, and 3) the generalizability of the LEAP to non-traditional applicants, as well as the susceptibility of the LEAP to socially desirable response bias.

Use of the LEAP in Conjunction with the AFOQT to Predict ROTC Training Performance

In keeping with Messick's (1989) argument that the validation of an assessment instrument involves hypothesizing and testing posited relationships, the following hypothesis was formulated:

Hypothesis 1: When used in conjunction with the AFOQT, the LEAP will increase the variance explained in a traditional training criterion over that obtained by the use of the AFOQT alone.

Testing Hypothesis 1. At the end of their summer encampments, ROTC cadets were rated by their supervisors on 10 field training criteria as specified by AFROTC Form 708, Cadet Field Training Performance Report (see Appendix C). The 10 criteria are:

1. *Adaptability to Military Training (AMT)*: cadet respects authority, adheres to standards and rules, exercises self-discipline, and functions effectively within the field training environment.

2. *Duty Performance (DtyP)*: cadet successfully completes assigned tasks in a timely manner and demonstrates sound judgment, imagination, self-discipline, and a willingness to perform these duties.

3. *Leadership/Followership (LdrF)*: cadet willingly accepts leadership responsibility, displays decisiveness and initiative in problem solving, and demonstrates interpersonal skills required to assist team members in task accomplishment.

4. *Adaptability to Stress (AStr)*: cadet displays an even temperament in a wide range of situations.

5. *Drill and Ceremonies (Drill* [AFROTC Form 204, Individual Drill Evaluation]*)*: cadet exhibits command voice, precision, bearing, alignment, and execution in drill and ceremony activities.

6. Human Relations (HumRl): cadet demonstrates empathy and sensitivity toward others and interpersonal skills that allow cadet to be an effective group member.

7. *Physical Fitness (PhFit)*: cadet performs satisfactorily in timed runs and physical fitness tests.

8. Communication Skills (Comm): cadet demonstrates ability to communicate in a clear and concise manner which is organized and grammatically correct, and demonstrates command of the language.

9. *Judgment and Decisions (JDec)*: cadet faces problems, appears in control, accepts and considers criticism, accepts own part in problem areas, and has ability to make decisions.

10. *Professional Qualities (PQual)*: cadet is cooperative, presents a professional military appearance, and demonstrates proper military bearing and presence, including proper use of military customs and courtesies.

A board consisting of the Field Training Camp Commander (FTCC), Commandant of Cadets (COC), and Flight Training Officers (FTOs) determined a class ranking for each cadet. The ranking was based on a composite of scores for each of the 10 attributes specified. The form requires that each

cadet be rated on a four-point scale corresponding to the following performance standards:

1 = Does not meet standards
2 = Meets standards but needs improvement
3 = Meets standards
4 = Exceeds standards

Once the field training performance ratings were received from AFROTC headquarters, two basic analyses were performed. First, descriptive statistics were computed to provide an overview of the ratings made on the 10 training performance scales. These results are presented in Table 16. It is apparent from the magnitude of the standard deviations and from a frequency analysis that the cadets were primarily rated "2" or "3." However, some cadets received extreme score ratings; so, the entire range of the four-point scale was used for all component scales. For the total scale, the obtained scores ranged from 10 to 36, covering 87% of the possible range.

Table 16. Descriptive Statistics for Field Training Performance Scores[a]

Component and Total Scores[b]	Mean	Standard Deviation
TOTAL	25.60	3.89
Adaptability to Military Training	2.78	.69
Duty Performance	2.84	.71
Leadership/Followership	2.73	.70
Adaptability to Stress	2.84	.64
Drill and Ceremonies	2.75	.72
Human Relations	2.82	.68
Physical Fitness	3.37	.60
Communication Skills	2.82	.61
Judgment and Decisions	2.66	.65
Professional Qualities	2.86	.71

[a]n = 506
[b]For each of the 10 training performance factors, scores range from a minimum of 1 to a maximum of 4; the total possible score ranges from 10 to 40.

Second, a principal components analysis was carried out to determine the interrelationships among the 10 component subscores and the total field training performance (FTP) score. A principal components analysis partitions all variance, weighting the variables to maximize the variance that goes into the first factor. Typically, that variance is substantial. In this instance, the variance accounted for by the first factor was particularly large, 81% of the total criterion variance.

Nine other factors collectively contributed an additional 19%, the second factor contributing only an additional 3%. Loadings for the first factor given in Table 17 are high, from .86 to .93.

Table 17. Loadings of Field Training Performance Scores on the First Principal Component[a]

Rating Scales	Loadings on First Principal Component
Adaptability to Military Training	.91
Duty Performance	.93
Leadership/Followership	.92
Adaptability to Stress	.92
Drill and Ceremonies	.86
Human Relations	.86
Physical Fitness	.86
Communication Skills	.89
Judgment and Decisions	.91
Professional Qualities	.91

[a] n = 506

The results given in Table 17 indicate a high degree of commonality across the 10 component criterion scores: Each of these criteria is measuring the same basic attribute. On that basis, it was appropriate to use the total field training performance score (FTP) for validating LEAP O-2D. Hence, all subsequent analyses involving the ROTC training performance criteria used that index only.

Validating LEAP O-2D against the FTP criterion was accomplished by computing correlations based on scores generated from both the rational and the Ordinal keys. The results are presented in Table 18.[3] Using the Ordinal key substantially increased--from .11 to .45--the validity coefficient achieved based on the rational key. Using the Nominal key, yielded an even higher validity coefficient, .61 (See Appendix D).

[3] For purposes of comparison, correlations were also run to determine outcomes when the Nominal and the Correspondence Analysis 1 scaling approaches were used. These results are presented in Appendix D.

Table 18. LEAP Scales Validated Against the Total Field Training Performance Score[a]

LEAP Scales	Rational Key	Ordinal Key
LEAP TOTAL	.11	.45***[b]
Transformational Leadership	.03	.21***
Transactional Leadership	.05	.04
Decision-Making Abilities	.05	.22***
Giving/Seeking Information	.07	.21***
Team Player Orientation	.10	.15**
Self-Sufficiency Orientation	.07	.25***
Physical Fitness Factors	.19	.35**
Institutional Commitment	.05	.14**
Persistence to Excellence	.11	.25***
Toleration of Adversity	.10	.10*
Socialized Power	.01	.22***
Retention Propensity	-.04	.08*

[a]n for the Rational Key = 328; for the Ordinal Key = 263
[b]Where p: * = <.05, ** = <.01, *** = <.001

Despite the fact that the Ordinal key did not achieve the highest correlation with the criterion, it was selected for use in subsequent analyses since, as indicated earlier, it represented the best compromise of *predictive efficiency* and adherence to the conceptual framework.

Finally, a stepwise regression analysis was performed using scores based on the Ordinal empirical key to determine which among the LEAP component scales were most robust in accounting for FTP criterion variance. When LEAP scale scores from the Ordinal empirical key were used to predict the FTP criterion, 27% of the criterion variance was explained. Results are presented in Table 19.

Having established that the Ordinally keyed LEAP scores were effective predictors of ROTC cadet field training performance as rated by supervisors, the next question was the crucial one: When used in conjunction with the Air Force Officer Qualifying Test (AFOQT; Berger, Gupta, Berger, & Skinner, 1988, 1990a), does the LEAP increase the variance explained in a traditional training performance criterion over that obtained by the AFOQT alone? To answer this question, AFOQT scores were obtained from AL/HR for the entire sample of ROTC cadets in 1991 summer encampments. Descriptive statistics for AFOQT scores achieved by this sample are presented in Table 20.

Means and standard deviations for this moderate sized sample appeared to be representative since they were consistent with those described by Skinner and Ree (1987). Next, regression analyses were performed in which the AFOQT was used to predict the Field Training Performance score (FTP). The results are summarized in Table 21.

Table 19. Regression Analysis Predicting the Total Field Training Performance Score, Based on the Ordinal Key[a]

Step	Variable Entered	Partial R^2	Cumulative Model R^2	F	p
1	PhysFit	.12	.12	44.18	.0001
2	PerExl	.04	.16	16.46	.0001
3	SocPwr	.03	.19	11.76	.0007
4	S-SOr	.03	.22	10.36	.0014
5	RetPrp	.02	.24	6.23	.0131
6	G/SInf	.01	.25	6.02	.0147
7	D-MAbl	.01	.26	4.55	.0336
8	TrfLdr	.01	.27	3.11	.0790

[a]n = 263
[b]PhysFit = Physical Fitness
PerExl = Persistence to Excellence
SocPwr = Socialized Power
S-SOr = Self-Sufficiency Orientation
RetPrp = Retention Propensity
G/SInf = Giving/Seeking Information
D-MAbl = Decision-Making Abilities
TrfLdr = Transformational Leadership

Table 20. Descriptive Statistics for AFOQT Scores for 1991 ROTC Summer Encampment Cadets

Scale Composite	Mean Raw Score	Standard Deviation	Minimum Score Obtained	Maximum Score Possible	Maximum Score Obtained
Pilot	121.56	23.23	48	205.0	184
NavTech	164.28	31.92	68	265.0	243
Academic	100.91	22.53	31	150.0	146
Verbal	49.11	12.03	18	75.0	74
Quantitative	51.80	12.97	13	75.0	75

[a]n = 721
Note: Minimum score possible for all scales was 0.

The AFOQT was a significant predictor of supervisor ratings (FTP), although it accounted for only 4% of the variance in the criterion. Because the AFOQT was used as a predictor with ROTC applicants, the range of scores is likely to be less than that for the total ROTC applicant pool. However, given that one-half or more of the ROTC applicants are selected, the restriction of range problem may not be as great as if that selection ratio were lower.

Table 21. Regression Analyses Using AFOQT, LEAP, and Combined Scores to Predict the Total Field Training Performance Score

Predictor	df	Sum of Squares	R^2	F	p
AFOQT	4	142.34	.04	2.52	.0400
12 LEAP Scales	12	1007.97	.27	7.57	.0001
AFOQT + 12 LEAP Scales	16	1114.23	.30	.42	.0001

To estimate the possible shrinkage effects in this situation, ordinarily a separate cross-validation sample would be used. However, Cronbach (1970, p. 430) suggests that when the selection ratio is high (i.e., the number of individuals selected is large relative to the number of applicants), and when the correlation between predictor(s) and criterion is low (as it is in this instance), the adjustment for the restriction in range will be minimal [see Cronbach (1970), Figure 13.6]. Thus, if the study had been based on the full range of scores, the obtained validity coefficient would not have been significantly larger than that obtained, r = .20.

Using the total LEAP score based on the Ordinal key, LEAP O-2D also proved to be a significant predictor of supervisor ratings. It accounted for 27% of the variance in the criterion, almost a seven-fold increment over the variance accounted for by the AFOQT.

Finally, a regression analysis was performed on a combined AFOQT model, in which the AFOQT composites were "forced" to be the first entrants into the equation, followed by the total LEAP score. Again, the combined model was used to predict the FTP criterion. The degree to which the combined model increased the percentage of variance accounted for in the criterion was the critical evidence needed to support or reject Hypothesis 1. This model accounted for 30% of the variance in the supervisor rating (FTP) criterion, an increment of 26% over that predicted by the AFOQT alone and 3% over that predicted by the LEAP alone.

Use of the LEAP for Diagnostic and Classification Purposes

As previously suggested, a second possible use of the LEAP would be for diagnostic or classification purposes. Such an application presumes that each component scale, in contrast to the LEAP total scale score, has sufficient utility to be used as a reliable indicant of the LEAP construct it was designed to

measure. Evidence supporting this use can be inferred from three sources: first, from an examination of the reliability coefficients for each of the component scales, second from an exploration of the internal structure of LEAP O-2D, and third, from an examination of the correlation of these component scales with a specially constructed set of criterion ratings. These ratings were designed to assess LEAP constructs using independent peer ratings. Evidence from these three sources will be presented to test the following two hypotheses:

Hypothesis 2: LEAP component scales will yield low, positive intercorrelations (i.e., r = .20 or less), with the exception of conceptually related scales, which will yield moderately positive intercorrelations (i.e., r = .35 or greater).

Hypothesis 3: Each of the component scales of the LEAP will correlate more highly with its counterpart dimension on the AFROTC Peer Rating Form than it will with any of the other sixteen dimensions of the form.

Testing Hypothesis 2. The moderate sized sample of LEAP O-2D respondents allowed assessment of both the reliability and the internal structure of the LEAP O-2D. Reliability was determined by a test-retest analysis and the internal structure of the measure was reflected by component scale intercorrelations.

Since reliability imposes a limitation upon the possible validity coefficients achieved, it was important to establish component scale reliabilities. As reported in Table 15, the LEAP O-2D demonstrated adequate (.71) overall test-retest reliability based on Ordinal key data. However, the corresponding component scales were marginal to low. With the exception of Persistence to Excellence, which achieved a coefficient of .78, six of the component scales had reliability coefficients in the .60's, two others were in the .50 to .60 range, and four others had coefficients in the .40 to .50 range. These component scale results require improvement before they can be confidently applied for diagnostic purposes. This is particularly true considering that the diagnostic use of the measure involves individual rather than grouped data.

On the basis of the initial conceptual model it was hypothesized that the successful Air Force officer possesses each of the attributes appraised by each of the component LEAP scales. Since these attributes are viewed as complementary, it was assumed that the more effective the Air Force Officer, the more strongly that officer would manifest each of the attributes by means of elevated scores on each of the LEAP component scales. Correspondingly, an intercorrelation matrix of those scales should reveal that the component scale scores are positively intercorrelated. In addition, because of the relative independence of these biographic scales, it was expected that the magnitude of the scale intercorrelations would be relatively modest, +.20 or less.

An exception to that generalization was expected in the case of conceptually linked component scales; that is, scales in which one was subsumed by another,

or was otherwise logically related to it. In that instance, the relationship was expected to be higher, +.35 or greater. An example of conceptually linked scales is Decision-Making Abilities and Transformational Leadership. Since the first of these scales is subsumed within the second, the link between those two component scales should be closer and the correlation coefficient higher than for either scale paired with non-conceptually linked scales. Six additional pairings were identified: Giving/Seeking Information and Transformational Leadership; both Decision-Making Abilities and Giving/Seeking Information paired with Transactional Leadership; Transformational Leadership and Team Player Orientation; and Socialized Power overlapping with Institutional Commitment and with Team Player Orientation. Setting the criterion levels at .20 and .35, respectively, was admittedly arbitrary, but seemed consistent with the above-mentioned considerations.

As shown by the data presented in Tables 22 and 23, the intercorrelations among the non-linked component scales, whether based on rational or Ordinal key data, yielded low positive coefficients (.20 or less) supportive of Hypothesis 2. These results also support the LEAP developers' intent to create a multi-trait selection device with substantial independence among the measure's 12 component scales.

However, with regard to the conceptually linked scales the evidence was not supportive of Hypothesis 2. Only one pair--Transformational Leadership and Decision-Making Abilities--met the standard posited on the Rational (.43) though not on the Ordinal key data (.16). Several other component scale pairs approached the specified coefficient: Institutional Commitment was correlated with Socialized Power on the Ordinal key at .30 and with Team Player Orientation at .28. However, the evidence does not consistently point in the direction postulated. That is, there is no consistent tendency for the magnitude of the coefficients to be higher for the conceptually linked component scales than for those not so identified.

In contrast to the erratic relationship between component scales, the data presented in Tables 22 and 23 show the much more consistent relationship between the total LEAP score and each of the components. Using the Ordinal key, the total LEAP scale is significantly correlated with all but one of the component scales, and with all of the scales when the rational key is used. Nonetheless, Tables 22 and 23 can be considered as multi-trait matrices, clarifying the interrelationships among all of the LEAP scales.

At this stage of development, the evidence only supports the use of the total, but not the component, LEAP scale scores. Hence, the LEAP is not yet suitable for use as a diagnostic device in identifying respondents' relative strengths and weaknesses. In addition to needed scale refinements, further research will be required to establish cutting scores for each of the scales before they may be used for diagnostic purposes.

Table 22. Intercorrelation Matrix for 12 LEAP Scales and Total LEAP Score, Based on the Rational Key[a,b]

	TOT LEAP	Trf Ldr	Trn Ldr	D-M Abl	G/S Inf	T-P Or	S-S Or	Phy Fit	Ins Com	Prs Excl	Tol Adv	Soc Pwr	Ret Prp
TotLEAP[c]													
TrfLdr	.80***												
TrnLdr	.26**	.11											
D-MAbl	.54***	.43***	.11*										
G/SInf	.49***	.33***	.13*	.08									
T-POr	.47***	.12*	.06	.13*	.08								
S-SOr	.46***	.28***	.08	.17**	.19**	.11*							
PhyFit	.35***	.20***	-.01	.16**	.20***	.15*	.23***						
InstCom	.41***	.11*	.09	.09	.04	.27***	.11	.06					
PrsExcl	.32***	.23***	.03	.13*	.14*	.03	.12*	.14*	-.01				
TolAdv	.46***	.20***	-.09	.11*	.08	.16***	.15*	.11*	.12*	.11*			
SocPwr	.54***	.31***	.08	.16**	.19**	.23***	.12*	.14*	.15*	.11*	.16*		
RetPrp	-.30***	-.29***	.05	-.28***	-.16**	-.02	-.15*	-.13*	.06	-.20***	-.10	-.10	

[a] n = 425

[b] Where p is: * = <.05, ** = <.01, *** = <.001

[c] TotLEAP = Total LEAP Score
TrfLdr = Transformational Leadership
TrnLdr = Transactional Leadership
D-MAbl = Decision-Making Abilities
G/SInf = Giving/Seeking Information
T-POr = Team Player Orientation
S-SOr = Self-Sufficiency Orientation
PhyFit = Physical Fitness Factors
InstCom = Institutional Commitment
PrsExcl = Persistence to Excellence
TolAdv = Toleration of Adversity
SocPwr = Socialized Power
RetPrp = Retention Propensity

Table 23. Intercorrelation Matrix for 12 LEAP Scales and Total LEAP Score, Based on the Ordinal Key[a,b]

	TOT LEAP	Trf Ldr	Trn Ldr	D-M Abl	G/S Inf	T-P Or	S-S Or	Phy Fit	Ins Com	Prs Excl	Tol Adv	Soc Pwr	Ret Prp
TotLEAP[c]													
TrfLdr	.68***												
TrnLdr	.19**	.08											
D-MAbl	.42***	.16**	.04										
G/SInf	.46***	.22***	.07	.16**									
T-POr	.43***	.08	.01	.12	.06								
S-SOr	.40***	.11	.10	.15*	.16*	.14*							
PhyFit	.42***	.18**	-.08	.09	.14*	.14*	.15						
InstCom	.45***	.12*	.03	.05	.06	.28***	.04	.09					
PrsExcl	.40***	.10	.09	.20***	.26***	-.01	.15*	.15*	-.03				
TolAdv	.12*	-.08	-.16**	.01	-.11	.13*	-.03	.07	.07	-.03			
SocPwr	.37***	.07	.01	.06	.13*	.16**	.08	.12	.30***	.03	.00		
RetPrp	.02	.04	-.09	-.11	-.10	-.03	.01	-.01.	.15*	-.04	-.15*	.06	

[a] n = 241

[b] Where p is: * = <.05, ** = <.01, *** = <.001

[c] TotLEAP = Total LEAP Score
TrfLdr = Transformational Leadership
TrnLdr = Transactional Leadership
D-MAbl = Decision-Making Abilities
G/SInf = Giving/Seeking Information
T-POr = Team Player Orientation
S-SOr = Self-Sufficiency Orientation
PhyFit = Physical Fitness Factors
InstCom = Institutional Commitment
PrsExcl = Persistence to Excellence
TolAdv = Toleration of Adversity
SocPwr = Socialized Power
RetPrp = Retention Propensity

Testing of Hypothesis 3. To augment the general ROTC Field Training Performance criteria, cadets responded to a 19-dimension, Air Force Peer Rating Form (AFPRF). This newly devised form was created specifically to tap the constructs the LEAP scales were intended to measure, as well as overall success and future potential as an Air Force officer. Unlike the Form 708, the AFPRF was intended to be completed by peers rather than by supervisors. Intentionally, these peers were flightmates who knew those well whom they were anonymously rating. The AFPRF is included as Appendix E.

The AFPRF was administered to the 225 ROTC cadets participating in the 1991 summer encampments at Lackland and Plattsburgh AFBs. The 19 dimensions of the peer rating form and the LEAP constructs to which they relate are shown in Table 24.

Descriptive statistics for the scores yielded by the ratings are summarized in Table 25. Note that dimensions #18 and #19 differ from the others. They represent overall ratings and are presented both separately and combined as indices of the total score. As reflected by the means and standard deviations, there was apparent reticence to use the upper end of the five-point peer rating scale, contributing to a restricted range in the ratings.

Intercorrelations among AFPRF criterion scales are presented in Table 26. Moderate intercorrelations are considered optimal. If the resulting correlation coefficients are too high, the dimensions may be faulted for redundancy; if they are too low, the lack of relevance of dimensions to each other may be questioned.

In fact, coefficients ranged widely from .29 to .86, clustering around .60. As expected, correlations between constructs thought to be unrelated, e.g., Decision-Making Abilities and Institutional Commitment (#8 and #14) tended to be lower (.34) than were correlations between constructs thought to be conceptually related. For example, two managerial constructs, Decision-Making Abilities and Giving/Seeking Information (#3 and #4), were correlated .86.

To gather evidence regarding Hypothesis 3, the AFPRF dimensions were correlated with their corresponding LEAP component scale scores. For comparative purposes, correlations were computed based on both rational and Ordinal keyed scores. The results obtained are reported in Tables 27 through 30 and in Appendix F.

Tables 27 through 30 report only partial data. Tables 27 and 29 reveal all significant intercorrelations between component scales and peer rating dimensions. Tables 28 and 30 are limited to the correlations between each LEAP component scale score and its AFPRF counterpart dimension, whether significant or not. Complete results are given in Appendix F.

Table 24. Corresponding Attributes Between LEAP O-2D Scales and 19 AFPRF Dimensions

Construct Measured	Corresponding Rating Dimension
Transformational Leadership (TrfLdr)	#1: When serving as leader, this cadet motivated others to go beyond their best previous levels of performance.
Transactional Leadership (TrnLdr)	#2: When serving as a leader, this cadet rewarded good performance and reprimanded poor performance of others.
Decision-Making Abilities (D-MAbl)	#3: This cadet could identify problems, analyze them, and then come up with effective solutions.
Giving/Seeking Information (G/SInf)	#4: By monitoring what was going on, this cadet gathered useful information, then shared it with others to better help the flight carry out its work.
Team-Player Orientation (T-POr)	#5: This cadet worked well with other flight members, drawing on each cadet's ideas, strengths, or resources to achieve the group's goals collaboratively.
Self-Sufficiency Orientation (S-SOr)	#6: This cadet worked effectively on his or her own, relying on his or her own judgment to make needed decisions.
Physical Fitness Factors (PhyFit)	#7: This cadet showed a concern for maintaining good health through willing participation in more than the required physical training.
Institutional Commitment (InstCom)	#8: This cadet willingly made personal sacrifices out of loyalty to the Air Force or out of commitment to its goals and values.
Persistence to Excellence (PrsExcl)	#9: This cadet worked hard on assigned duties and was not satisfied until the best possible performance was achieved.
Toleration of Adversity (TolAdv)	#10: This cadet worked hard at all duties or tasks despite any adversity or frustration experienced.

Table 24. (Concluded)

Construct Measured	Corresponding Rating Dimension
Socialized Power (SocPwr)	#11: This cadet listened to, advised, and supported others.
Socialized Power (SocPwr)	#12: This cadet encouraged others to take the work of the flight more seriously and to make a stronger commitment to the achievement of its goals.
Transformational Leadership (Charisma) (TrfLdr)	#13: This cadet inspired others and gained support for his/her suggestions and ideas.
Decision-Making Abilities (Problem Solving) (D-MAbl)	#14: This cadet found new and creative ways to solve problems or complete tasks.
Transformational Leadership (Individualized Consideration) (TrfLdr)	#15: In a leadership position, this cadet considered the needs and abilities of others when assigning tasks or duties.
Transformational Leadership Intellectual Stimulation) (TrfLdr)	#16: This cadet motivated others to act by raising challenging problems or questions for them to solve. This cadet helped others find new ways to think and to handle tasks or assignments.
Decision-Making Abilities (Planning and Organizing) (D-MAbl)	#17: This cadet planned or carried out tasks in an organized fashion.
Overall Successful Performance (Total LEAP)	#18: This cadet demonstrated qualities that resulted in a high degree of success during this encampment.
Future Potential (Total LEAP)	#19: This cadet demonstrated qualities that show the potential for becoming an outstanding future Air Force officer.

Table 25. Descriptive Statistics for Peer Rating Dimensions[a]

AFPRF Dimensions	LEAP Scales[b]	Mean	Minimum Standard Deviation	Maximum Obtained Score	Maximum Possible Score	Obtained Score
#1,13,15,16	TrfLdr[c]	8.21	2.27	2.63	20.00	13.60
#2	TrnLdr	2.11	0.64	0.56	5.00	3.72
#3,14,17	D-MAbl	6.44	1.79	1.80	15.00	10.90
#4	G/SInf	2.24	0.68	0.42	5.00	3.92
#5	T-POr	2.03	0.68	0.36	5.00	3.80
#6	S-SOr	2.19	0.66	0.40	5.00	3.64
#7	PhyFit	2.12	0.71	.00	5.00	3.75
#8	InstCom	2.40	0.64	0.64	5.00	3.89
#9	PrsExcl	2.23	0.64	0.73	5.00	3.70
#10	TolAdv	2.44	0.59	0.54	5.00	3.75
#11,12	SocPwr	4.44	1.14	1.83	10.00	7.18
TOT 1-17	TotLEAP	35.94	8.94	12.75	85.00	57.40
#18	TotLEAP	1.98	0.70	0.30	5.00	3.80
#19	TotLEAP	2.37	0.60	0.80	5.00	3.90
TOT 18/19	TotLEAP	4.35	1.25	1.50	10.00	7.60

[a] n = 225
[b] Refers to corresponding LEAP component scale
[c] TrfLdr = Transformational Leadership
TrnLdr = Transactional Leadership
D-MAbl = Decision-Making Abilities
G/SInf = Giving/Seeking Information
T-POr = Team Player Orientation
S-SOr = Self-Sufficiency Orientation
PhyFit = Physical Fitness Factors
InstCom = Institutional Commitment
PrsExcl = Persistence to Excellence
TolAdv = Toleration of Adversity
SocPwr = Socialized Power
TotLEAP = Total LEAP Score
Note: Minimum possible score for all dimensions is 0.

Table 26. Intercorrelations Among ROTC Peer Rating Dimensions[a]

	1[b]	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	1-17 TOT	18	19	18/19 TOT
1																					
2	.66																				
3	.69	.70																			
4	.65	.70	.86																		
5	.70	.59	.75	.75																	
6	.69	.56	.68	.58	.66																
7	.63	.62	.61	.60	.60	.52															
8	.64	.45	.62	.55	.69	.61	.60														
9	.70	.72	.77	.70	.69	.72	.59	.72													
10	.67	.61	.74	.71	.73	.66	.63	.73	.80												
11	.55	.41	.57	.61	.71	.47	.43	.54	.52	.65											
12	.59	.57	.64	.65	.53	.52	.44	.47	.66	.64	.45										
13	.65	.76	.74	.79	.68	.52	.64	.46	.62	.65	.55	.63									
14	.44	.64	.66	.67	.56	.36	.54	.34	.49	.55	.39	.54	.78								
15	.54	.52	.68	.66	.68	.55	.50	.61	.64	.57	.53	.47	.57	.44							
16	.49	.65	.64	.68	.49	.42	.49	.29	.45	.50	.38	.58	.73	.66	.51						
17	.64	.49	.75	.69	.73	.72	.50	.70	.77	.74	.57	.55	.49	.39	.64	.40					
TOT 1-17	.81	.75	.88	.85	.84	.78	.73	.73	.84	.86	.69	.72	.81	.66	.73	.68	.80				
18	.80	.53	.76	.71	.76	.72	.62	.71	.76	.75	.64	.62	.63	.44	.61	.44	.81	.86			
19	.73	.56	.70	.67	.70	.75	.56	.59	.72	.73	.60	.63	.66	.49	.60	.49	.71	.83	.82		
TOT 18/19	.76	.54	.73	.69	.73	.73	.59	.65	.74	.74	.62	.62	.64	.47	.60	.46	.75	.86	.96	.94	

[a] n = 225

[b] #1 = motivates others to go beyond previous best
#2 = rewards good, reprimands poor performance
#3 = identifies problems, determines solutions
#4 = gathers information, shares with others
#5 = works well with others, draws on their talents
#6 = works well on own, relies on own judgment
#7 = maintains good health through physical training
#8 = makes personal sacrifices for the Air Force
#9 = works hard until best performance is achieved
#10 = works hard despite adversities
#11 = listens to, advises, and supports others
#12 = encourages others to take work seriously
#13 = inspires others and gains their support
#14 = finds new ways to solve problems
#15 = assigns duties based on needs and abilities
#16 = motivates others by presenting challenging tasks
#17 = plans/carries out tasks in organized manner
TOT 1-17 = composite of dimensions numbered 1-17
#18 = demonstrates successful qualities in present
#19 = shows qualities for success in future
TOT18/19 = composite of dimensions #18 and #19

Table 27. Significant Intercorrelations[a] Among LEAP O-2D Component Scales and Peer Ratings, Based on the Rational Key[b]

	19 Peer Rating Dimensions																							
LEAP Scales	**#1[c]**	**#2**	**#3**	**#4**	**#5**	**#6**	**#7**	**#8**	**#9**	**#10**	**#11**	**#12**	**#13**	**#14**	**#15**	**#16**	**#17**	**TOT 1-17**	**#18**	**#19**	**TOT 18/ 19**	**TOT Trf Ldr**	**TOT D-M Abl**	**TOT Soc Pwr**
TotLEAP[d]	.21	.16	.19	.16	.12	.18	.13	.15	.18	.18	.10	.15	.12	.16	.05	.11	.13	.21	.17	.15	.17	.16	.18	.14
TrfLdr	.23	.15.	.20	.18	.15	.17	.12	.13	.13	.17	.15	.17	.15	.10	.10	.13	.10	.21	.20	.21	.21	.20	.16	.18
Chrs	.15	.11	.18	.17	.14	.10	.10	.10		.16	.12	.12	.15	.11		.14		.17	.13	.14	.14	.17	.13	.14
TrnLdr																								
D-MAbl	.12	.17	.10	.12	.12	.13	.25	.11	.11			.10			.18	.14	.10	.18	.15	.19	.17	.15	.10	
G/SInf	.16	.13	.24	.26	.22	.18	.16	.19	.18	.29	.35	.14	.16	.21			.33	.25	.19	.18	.19	.13	.30	.29
T-POr	.20					.10	.14	.17	.16			.12						.11	.14		.12		.10	
S-SOr																								
PhyFit	.25	.22	.26	.26	.27	.22	.19	.19	.25	.34	.24	.22	.21	.28	.16	.22	.31	.31	.29	.25	.29	.26	.34	.27
InstCom		-.11			-.13		-.11	.10			-.14		-.17		-.19	-.11							-.13	
PrsExcl		.22	.10	.16	.16		.10			.16			.20	.19	.10	.21	.14	.16				.18	.17	
TolAdv																								
SocPwr		.10				.13			.12			.12												

[a] .10 or higher
[b] n = 127
[c] #1 = motivates others to go beyond previous best
#2 = rewards good, reprimands poor performance
#3 = identifies problems, determines solutions
#4 = gathers information, shares with others
#5 = works well with others, draws on their talents
#6 = works well on own, relies on own judgment
#7 = maintains good health through physical training
#8 = makes personal sacrifices for the Air Force
#9 = works hard until best performance is achieved
#10 = works hard despite adversities
#11 = listens to, advises, and supports others
#12 = encourages others to take work seriously
#13 = inspires others and gains their support
#14 = finds new ways to solve problems
#15 = assigns duties based on needs and abilities
#16 = motivates others by presenting challenging tasks
#17 = plans/carries out tasks in organized manner

TOT 1-17 = composite of dimensions 1-17
#18 = demonstrates successful qualities in present
#19 = shows qualities for success in future
TOT 18/19 = composite of dimensions #18 and #19
TOT TrfLdr = composite of #1, #13, #15 and #16
TOT D-MAbl = composite of #14 and #17
TOT SocPwr = composite of #11 and #12
[d] TotLEAP = Total LEAP Score
TrfLdr = Transformational Leadership
Chrs = Charisma
TrnLdr = Transactional Leadership
D-MAbl = Decision-Making Abilities
G/SInf = Giving/Seeking Information
T-POr = Team Player Orientation
S-SOr = Self-Sufficiency Orientation
PhyFit = Physical Fitness Factors
InstCom = Institutional Commitment
PrsExcl = Persistence to Excellence
TolAdv = Toleration of Adversity
SocPwr = Socialized Power

Table 28. Intercorrelations Among LEAP O-2D Component Scales and Corresponding[a] Peer Ratings, Based on the Rational Key[b]

	19 Peer Rating Dimensions																							
LEAP Scales	**#1[c]**	**#2**	**#3**	**#4**	**#5**	**#6**	**#7**	**#8**	**#9**	**#10**	**#11**	**#12**	**#13**	**#14**	**#15**	**#16**	**#17**	**TOT 1-17**	**#18**	**#19**	**TOT 18/ 19**	**TOT Trf Ldr**	**TOT D-M Abl**	**TOT Soc Pwr**
TotLEAP[c]																								
TrfLdr	.23												.15		.10	.13						.20		
Chrs													.15											
TrnLdr		.07																						
D-MAbl			.10																					
G/SInf				.26										.05			.10						.10	
T-POr					.07																			
S-SOr						.02																		
PhyFit							.19																	
InstCom								.10																
PrsExcl									.06															
TolAdv										-.04														
SocPwr											-.08	.12												.02

[a]only correlations between scales and peer ratings presumed to measure the same construct are presented

[b]n = 127

[c]#1 = motivates others to go beyond previous best
#2 = rewards good, reprimands poor performance
#3 = identifies problems, determines solutions
#4 = gathers information, shares with others
#5 = works well with others, draws on their talents
#6 = works well on own, relies on own judgment
#7 = maintains good health through physical training
#8 = makes personal sacrifices for the Air Force
#9 = works hard until best performance is achieved
#10 = works hard despite adversities
#11 = listens to, advises, and supports others
#12 = encourages others to take work seriously
#13 = inspires others and gains their support
#14 = finds new ways to solve problems
#15 = assigns duties based on needs and abilities
#16 = motivates others by presenting challenging tasks
#17 = plans/carries out tasks in organized manner

TOT 1-17 = composite of dimensions 1-17
#18 = demonstrates successful qualities in present
#19 = shows qualities for success in future
TOT 18/19 = composite of dimensions #18 and #19
TOT TrfLdr = composite of #1, #13, #15 and #16
TOT D-MAbl = composite of #14 and #17
TOT SocPwr = composite of #11 and #12
[c]TotLEAP = Total LEAP Score
TrfLdr = Transformational Leadership
Chrs = Charisma
TrnLdr = Transactional Leadership
D-MAbl = Decision-Making Abilities
G/SInf = Giving/Seeking Information
T-POr = Team Player Orientation
S-SOr = Self-Sufficiency Orientation
PhyFit = Physical Fitness Factors
InstCom = Institutional Commitment
PrsExcl = Persistence to Excellence
TolAdv = Toleration of Adversity
SocPwr = Socialized Power

Table 29. Significant Intercorrelations[a] Among LEAP O-2D Component Scales and Peer Ratings, Based on the Ordinal Key[b]

	19 Peer Rating Dimensions																							
LEAP Scales	**1[c]**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**TOT 1-17**	**18**	**19**	**TOT 18/ 19**	**TOT Trf Ldr**	**TOT D-M Abl**	**TOT Soc Pwr**
TotLEAP[d]	.11	.10	.28	.36	.21	.23	.17	.25	.18	.24	.35	.14		.18			.26	.22	.25	.25	.27	.10	.27	.29
TrfLdr		.12	.18		.10												.14			.13	.10		.11	
Chrs			.14	.19						.15	.17			.11		.14					.11		.13	.10
TrnLdr		.10		.11							.13	.21	.14	.16	.11	.12		.13		.17		.12		.19
D-MAbl	-.18					.18						-.26				-.15						-.12		-.13
G/SInf	.10	.12	.22	.26	.21	.28	.25	.27	.12	.26	.29	.13	.10	.20			.28	.24	.19	.22	.22		.25	.25
T-POr					.13			.14											.13		.12			
S-SOr	.14	.11	.24	.26	.28	.36	.25	.27	.19	.19	.28		.25	.29	.18	.17	.24	.30	.23	.22	.25	.20	.27	.21
PhyFit	.17		.22	.20			.10	.17	.20	.24	.46	.13		.14	.16		.18	.15	.19	.14	.18	.11	.20	.35
InstCom	.22	.11	.23	.22				.15	.16	.11		.21		.18	.13		.14	.11	.17		.14	.12	.19	.14
PrsExcl			.20	.13	.17	.15				.19	.17	.10	.13	.14			.12		.13		.11		.16	.15
TolAdv		-.12			-.11						.10	-.11		-.17			-.13						-.11	
SocPwr		.11	.18					.13	.10	.12		.20	.13				.10							.19

[a] .10 or higher
[b] n = 127
[c] #1 = motivates others to go beyond previous best
#2 = rewards good, reprimands poor performance
#3 = identifies problems, determines solutions
#4 = gathers information, shares with others
#5 = works well with others, draws on their talents
#6 = works well on own, relies on own judgment
#7 = maintains good health through physical training
#8 = makes personal sacrifices for the Air Force
#9 = works hard until best performance is achieved
#10 = works hard despite adversities
#11 = listens to, advises, and supports others
#12 = encourages others to take work seriously
#13 = inspires others and gains their support
#14 = finds new ways to solve problems
#15 = assigns duties based on needs and abilities
#16 = motivates others by presenting challenging tasks
#17 = plans/carries out tasks in organized manner

TOT 1-17 = composite of dimensions 1-17
#18 = demonstrates successful qualities in present
#19 = shows qualities for success in future
TOT 18/19 = composite of dimensions #18 and #19
TOT TrfLdr = composite of #1, #13, #15 and #16
TOT D-MAbl = composite of #14 and #17
TOT SocPwr = composite of #11 and #12
[d] TotLEAP = Total LEAP Score
TrfLdr = Transformational Leadership
Chrs = Charisma
TrnLdr = Transactional Leadership
D-MAbl = Decision-Making Abilities
G/SInf = Giving/Seeking Information
T-POr = Team Player Orientation
S-SOr = Self-Sufficiency Orientation
PhyFit = Physical Fitness Factors
InstCom = Institutional Commitment
PrsExcl = Persistence to Excellence
TolAdv = Toleration of Adversity
SocPwr = Socialized Power

Table 30. Intercorrelations Among LEAP O-2D Component Scales and Corresponding[a] Peer Ratings, Based on the Ordinal Key[b]

	19 Peer Rating Dimensions																						
LEAP Scales	**1[c]**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**TOT 1-17**	**18**	**19**	**TOT 18/ 19**	**TOT Trf Ldr**	**TOT D-M Abl**	**TOT Soc Pwr**
Tot LEAP[c]																								
TrfLdr	.03												-.08		-.06	.01							-.02	
Chrs													-.08											
TrnLdr		.10																						
D-MAbl			-.06											-.02			-.03							-.04
G/SInf				.26																				
T-POr					.13																			
S-SOr						.36																		
PhyFit							.10																	
InstCom								.15																
PrsExcl									.09															
TolAdv										.02														
SocPwr											.20	.13												.19

[a]only correlations between scales and peer ratings presumed to measure the same constructs are presented

[b]n = 127

[c]#1 = motivates others to go beyond previous best
#2 = rewards good, reprimands poor performance
#3 = identifies problems, determines solutions
#4 = gathers information, shares with others
#5 = works well with others, draws on their talents
#6 = works well on own, relies on own judgment
#7 = maintains good health through physical training
#8 = makes personal sacrifices for the Air Force
#9 = works hard until best performance is achieved
#10 = works hard despite adversities
#11 = listens to, advises, and supports others
#12 = encourages others to take work seriously
#13 = inspires others and gains their support
#14 = finds new ways to solve problems
#15 = assigns duties based on needs and abilities
#16 = motivates others by presenting challenging tasks
#17 = plans/carries out tasks in organized manner

TOT 1-17 = composite of dimensions 1-17
#18 = demonstrates successful qualities in present
#19 = shows qualities for success in future
TOT 18/19 = composite of dimensions #18 and #19
TOT TrfLdr = composite of #1, #13, #15 and #16
TOT D-MAbl = composite of #14 and #17
TOT SocPwr = composite of #11 and #12
[c]TotLEAP = Total LEAP Score
TrfLdr = Transformational Leadership
Chrs = Charisma
TrnLdr = Transactional Leadership
D-MAbl = Decision-Making Abilities
G/SInf = Giving/Seeking Information
T-POr = Team Player Orientation
S-SOr = Self-Sufficiency Orientation
PhyFit = Physical Fitness Factors
InstCom = Institutional Commitment
PrsExcl = Persistence to Excellence
TolAdv = Toleration of Adversity
SocPwr = Socialized Power

As shown in Table 29, the Ordinal key derived total LEAP score yielded a significant correlation of .22 with the sum of the 17 component peer rating dimensions and correlated .27 with the global #18/#19 criterion. However, the LEAP component scale data presented in Table 29 reveal that the component scale scores varied widely in the degree to which they were predictive of their corresponding peer rating dimensions. Three scales--Giving/Seeking Information, Self-Sufficiency Orientation, and Socialized Power--produced significant correlations with their corresponding dimensions (#4, #6, #11, respectively), ranging from .20 to .36. The correlations for all of the remaining scales with their counterpart dimensions were generally lower, yielding non-significant validity coefficients ranging from .00 to .10.

On both the rational and Ordinal keyed scores presented in Tables 27 and 29, the Physical Fitness Factors dimension (PhyFit) was significantly correlated with virtually all other peer rating dimensions. This result calls the appropriateness of the peer ratings into question as a suitable set of criteria for this population. These data suggest that the physical fitness level of a cadet is driving ratings made on all other dimensions. That is, a cadet with a high level of physical fitness seems to earn high peer ratings on all other dimensions.

Thus, the intercorrelation results do not support Hypothesis 3. Although the total LEAP score correlates significantly with total peer rating score, most of the LEAP component scales do not correlate significantly with their corresponding peer rating. ROTC cadets may not yet have the maturity to rate their classmates reliably, or the experimental AFPRF at this stage of development was not suited to the purpose for which it was applied. It is also possible that these outcomes could be attributed to weaknesses in the LEAP instrument itself, though the performance of the LEAP as an effective predictor against the Field Training Performance (FTP) criterion lessens the likelihood of that interpretation.

Generalizability of the LEAP to Multiple Applicant Groups

Unlike the other two portions of this validity section, this third portion does not address a direct application. Rather, it provides a procedural check of the applicability of the LEAP to all respondents. This section addresses the following hypotheses:

Hypothesis 4: There will be no evidence of participant response bias on the basis of gender, ethnicity, or socioeconomic status (SES) level.

Hypothesis 5: There will be no evidence of participants responding to questions in a socially desirable manner rather than a manner reflecting their actual experience.

Testing Hypothesis 4. Tests to determine bias for gender, ethnicity, or socioeconomic status were performed. The sample was divided into subgroups:

(1) male versus female; (2) white versus non-white; and (3) total family annual income greater than and less than $40,000. The mean scale and total LEAP scores for the subgroups are compared and presented in Table 31.

As can be seen in Table 31, differences between subgroup means were very small. The overall results provide evidence in support of Hypothesis 4: there are no systematic differences in the responses of participants on the basis of gender, ethnicity, or SES.

Testing Hypothesis 5. The accuracy of biodata has been documented in the literature by various authors, including Doll (1971) and Shaffer, Saunders, and Owens (1986). However, they have identified several factors that reduce response accuracy: (a) items which inquire into respondents' feelings and perceptions rather than behaviors or past events; (b) lack of a specific time referent; (c) items with continuous alternatives ranging from "Always" to "Never"; and (d) intentional distortion in order to represent the respondent in a more favorable light.

One of the most direct and effective methods for dealing with inaccurate responses is to use a Faking Detection scale. As Trent (1987) pointed out, military applicants share with other individuals in selection situations a desire to present themselves as favorably as possible. Thus, a Faking Detection scale was included in LEAP O-2D. These items were selected from the best of an original pool of 30 items pretested on a group of 20 ROTC cadet volunteers from among a group of 146 attending a summer encampment at Lackland AFB. First, the cadets were asked to identify which were Faking Detection items. Next, they were given a list of correctly identified Faking Detection items and were asked to rate each item on a five-point scale with regard to: (a) how detectable (or obvious) each item was; (b) how relevant each item was to the section in which it was embedded; (c) how socially desirable or threatening each item was; and (d) how easy or difficult each item was to understand. On the basis of feedback received, the Faking Detection scale was reduced to 15 items. Because there was only one "correct" response alternative for each item, there was no need to transform scores on the basis of an empirical key. Therefore, Faking Detection scores were based on a rational key.

The new 15-item Faking Detection scale was incorporated into an abbreviated LEAP (62 items) and administered to the entire group of summer encampment ROTC cadets at Lackland AFB. They completed the instrument three times under the following response conditions: fake good, fake bad, answer honestly. Descriptive statistics of the results obtained are presented in Table 32. As intended, the "fake good" condition generated the highest obtained mean Faking Detection score (8.90), while the "fake bad" condition generated the lowest (.38). The "honest response" condition resulted in an intermediate mean Faking Detection score (3.07).

Faking Detection items which revealed a large discrepancy between fake good and fake bad conditions, and which departed markedly from the honest responses, were included in a final 12-item scale embedded into LEAP O-2D.

Table 31. Gender, Ethnicity, and SES Differences in LEAP O-2D Mean Ordinal Scores[a]

	Means and Standard Deviations					
	Gender		Ethnicity		SES	
	Male	Female	White	NonWhite	Low	High
LEAP Scales	M SD	M SD	M SD	M SD	M SD	M SD
TotLEAP	302.7	302.7	302.7	302.57	302.7	302.6
	.74	.76	.75	.75	.77	.68
TrfLdr[b]	57.6	57.6	57.7	57.66	57.69	57.60
	.26	.30	.26	.31	.25	.20
Chrs	37.4	37.4	37.5	37.47	37.52	37.42
	.18	.20	.21	.19	.17	.18
TrnLdr	20.2	20.2	20.18	20.19	20.19	20.17
	.10	.10	.08	.08	.08	.07
D-MAbl	20.1	20.2	20.2	20.15	20.21	20.18
	.16	.11	.14	.14	.12	.10
G/SInf	20.1	20.1	20.2	20.16	20.16	20.16
	.14	.16	.14	.16	.16	.17
T-POr	20.1	20.2	20.2	20.15	20.16	20.15
	.11	.11	.13	.12	.16	.14
S-SOr	20.1	20.2	20.2	20.15	20.16	20.15
	.15	.12	.14	.13	.16	.15
PhyFit	25.9	25.9	26.0	25.88	25.95	25.9
	.27	.29	.26	.24	.24	.30
InstCom	20.1	20.2	20.2	20.19	20.2	20.2
	.13	.14	.15	.15	.15	.16
PrsExcl	20.1	20.2	20.2	20.17	20.17	20.18
	.18	.16	.16	.15	.17	.17
TolAdv	23.0	23.0	23.1	23.06	23.07	23.06
	.09	.08	.05	.05	.05	.05
SocPwr	34.5	34.5	34.6	34.58	34.57	34.6
	.12	.11	.14	.13	.13	.13
RetPrp	20.0	20.2	20.2	20.19	20.19	20.21
	.21	.14	.11	.10	.09	.16

[a] n = 259

[b] TrfLdr = Transformational Leadership
Chrs = Charisma
TrnLdr = Transactional Leadership
D-MAbl = Decision-Making Abilities
G/SInf = Giving/Seeking Information
T-POr = Team Player Orientation
S-SOr = Self-Sufficiency Orientation
PhyFit = Physical Fitness Factors
InstCom = Institutional Commitment
PrsExcl = Persistence to Excellence
SocPwr = Socialized Power
RetPrp = Retention Propensity

Table 32. Faking Detection Scale Scores Under Three Conditions[a]

Response Condition	Mean	Standard Deviation	Range
Fake Good	8.90	2.73	0-12
Fake Bad	.38	1.12	0-11
Honest Response	3.07	2.29	0-10

[a] n = 146

After LEAP O-2D was administered to cadets at their summer encampments, the responses to the Faking Detection scale were analyzed separately from those of other LEAP scales. Based on the responses of 425 ROTC cadets, the scale yielded a mean score of 4.71 and a standard deviation of 1.70. An alternative means of describing the response distribution of the Faking Detection scale is a bar graph presented in Figure 1. The figure shows a near normal distribution of scores with a slight skewness toward the upper end of the scale.

The Faking Detection scores were used to evaluate the degree to which "gaming" of the instrument occurred. The investigators decided to use an arbitrary score of 1.75 standard deviations or more above the mean for the "honest response" condition as a cutting score indicating possible gaming. Using this cut-off point, respondents achieving scores of 7.0 or above (7.08 = 1 3/4 S.D.'s above the mean) would be suspect as "gamers." As seen in Figure 1, this decision rule would isolate only 45 out of 425 respondents (10.59%). Of these, 29 had scores of 7, 13 had scores of 8, 2 had scores of 9, and one had a score of 10. No respondents achieved scores of 11 or 12, the total number possible. The relatively small proportion of respondents scoring high on the Faking Detection scale suggests that no substantial "gaming" of the instrument was taking place.

A second test for possible effects of social desirability was performed. Mean component and total LEAP scale scores were compared in an extreme group analysis. Participants having a Faking Detection score of 7.0 or higher (n = 63) were compared to those with scores of 2.0 or below (n = 44). The results of this analysis are presented in Table 33.

The mean scores for the two extreme groups differed significantly on 10 of the 12 component scales (13 if Charisma is included separately). The difference in component scale means ranged from a differential of .09 to one of 1.66, with a mean difference value of .86. Given that each of the scales extends from at least 0 to 7, and that the average difference was less than 1.0, the extent of the distortion effect attributable to intentional or unintentional gaming of the LEAP may not be of much practical significance.

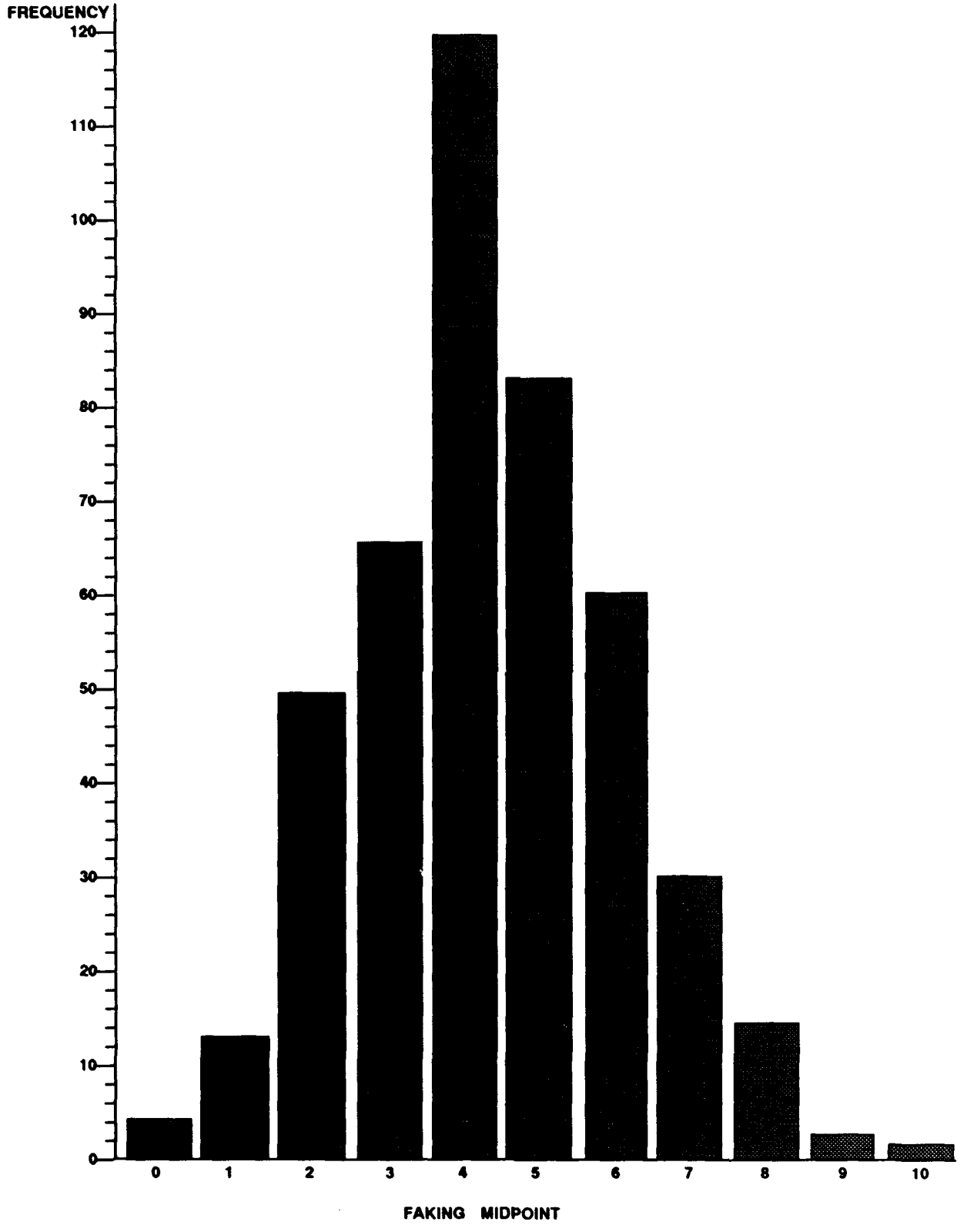


Figure 1. Distribution of Faking Scale Response Frequencies.

Table 33. Analysis of Variance and Related Comparisons for Highest and Lowest Scorers on the Faking Detection Scale[a]

Scale	Mean		Standard Deviation		F	p	Minimum Score Obtained		Maximum Score Possible	Maximum Score Obtained	
	High Group	Low Group	High Group	Low Group			High Group	Low Group		High Group	Low Group
TrfLdr[b]	10.76	9.10	2.06	2.70	11.77	.05	6.81	3.59	23.00	15.23	14.90
Chrs	6.74	5.89	1.38	1.89	6.48	.05	3.99	1.83	16.00	10.23	10.58
TrnLdr	4.06	4.20	1.17	1.07	.44	n.s.	1.16	1.83	8.00	5.91	6.41
D-MAbl	4.81	4.19	.98	.99	10.06	.05	2.81	1.16	7.00	7.00	6.16
G-SInf	4.58	4.79	.85	.81	23.86	.01	2.51	1.73	7.00	6.66	6.21
T-POr	4.66	3.30	1.02	1.10	41.95	.01	2.00	.50	7.00	7.00	5.33
S-SOr	4.50	3.83	.86	.92	14.49	.05	2.74	2.00	7.00	5.91	5.58
PhysFit	5.93	4.73	1.25	1.17	25.46	.01	3.35	2.01	9.00	8.25	6.80
InstCom	4.86	4.08	.94	.84	20.19	.01	2.64	2.15	7.00	6.67	5.98
PrsExcl	4.10	4.01	1.01	.98	24	n.s.	2.25	2.00	7.00	6.75	6.20
TolAdv	5.22	3.98	1.02	1.10	34.28	.001	2.99	1.50	8.00	8.00	6.83
SocPwr	6.69	5.16	1.35	1.39	32.39	.001	3.50	1.91	12.00	9.75	7.66
RetPrp	3.25	3.52	.98	1.07	1.81	n.s.	.50	.75	7.00	5.25	5.75

[a]n for High Group (score of 7 and above) = 44; Low Group (score of 2 and below) = 63

[b]TotLEAP = Total LEAP Score
TrfLdr = Transformational Leadership
Chrs = Charisma
TrnLdr = Transactional Leadership
D-MAbl = Decision-Making Abilities
G/SInf = Giving/Seeking Information
T-POr = Team Player Orientation
S-SOr = Self-Sufficiency Orientation
PhyFit = Physical Fitness Factors
InstCom = Institutional Commitment
PrsExcl = Persistence to Excellence
TolAdv = Toleration of Adversity
SocPwr = Socialized Power
RetPrp = Retention Propensity

Taken in combination with the prior analysis, the extent to which LEAP O-2D participants distorted their responses in a socially desirable manner is distinctly limited. Thus, Hypothesis 5 is essentially supported.

It must not be forgotten, however, that all LEAP O-2D respondents in this study had been admitted into ROTC. They do not constitute a sample of the actual LEAP target populations: ROTC or OTS applicants. Hence, the true *proof of the propensity to game the LEAP must await field testing of these* populations.

CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS

Conclusions

Five hypotheses were tested in attempting to gather evidence about the validity of the LEAP O-2D (ROTC). Three of those were fully supported. In support of the study's major focus, it was found that:

1. Consistent with Hypothesis 1, when used in conjunction with the AFOQT, the LEAP increased the R^2 for the AFOQT from .04 to .30.

In pursuit of supplemental objectives, it was ascertained that:

2. Hypothesis 2 was only partially supported. Although the intercorrelations among the non-conceptually linked LEAP component scales were of the low, positive magnitude hypothesized, the intercorrelations for the conceptually linked scales failed to reach the predicted magnitude of .35. The heterogeneous nature of these biodata scales probably made difficult the achievement of that minimum coefficient.

3. Hypothesis 3 was not supported. Intercorrelations between intercorrelations with the 19-dimension peer rating measure did not reveal the one-to-one correspondence hoped for between the LEAP component scales and their counterparts on the AFPRF. Further development on the empirical key and on the peer rating form is required.

4. Consistent with Hypothesis 4, the LEAP instrument does not appear to be biased based on gender, ethnicity, or SES. Responses of males versus females, whites versus non-whites, and respondents from families having high versus low annual incomes were not found to differ.

5. Consistent with Hypothesis 5, there was only limited evidence of faking proneness among ROTC respondents. Only l0% of respondents achieved scores on the Faking Detection scale higher than 7.0 based on a 12-point scale. Furthermore, although significant mean differences were found on 10 of 12 component scales, those differences appeared too small to have practical significance.

Implications of the Study

This study showed that the Air Force could effectively supplement the AFOQT in its selection of officers by using the LEAP. The LEAP increased the predictive power of the AFOQT substantially. The data showed that the LEAP contributed substantially to the predictive power of the AFOQT alone when correlated with a typical training performance criterion. Moreover, this outcome was achieved despite less than ideal conditions of survey administration. For example, the LEAP was administered by personnel not previously familiar with the instrument and given to a limited sample of respondents who had no particular motivation for the task. Finally, even at this relatively early stage in its development, the LEAP has demonstrated only limited incidence of response bias.

The outcomes above argue for the promise of the LEAP in the selection of ROTC cadets and perhaps OTS cadets. They also support continued refinement. What is particularly needed is evidence of the utility of the LEAP when administered to an actual target population--ROTC and OTS applicants. Better criteria against which to validate the LEAP are also greatly needed.

Recommendations for Further Research

Next steps in the continued refinement of the LEAP would involve further testing of the measure itself, together with a validation effort using Air Force ROTC applicants.

Steps recommended in the refinement of the LEAP are:

1. Revise LEAP O-2D (ROTC) using item analyses, with particular attention to items in the three new component LEAP scales. Use item weights for each response alternative generated by the empirical keys as another basis for refinement. For example, current rationally derived, item scoring weights could be compared to empirically derived weights to determine if they are functioning as intended.

2. Reexamine the conceptual framework of the LEAP in the light of findings from field testing experience to date.

3. Increase the size of the item pool by constructing at least 15 additional items for each of the LEAP component scales as recommended by the Laboratory Advisory Group (LAG).

4. Field test newly devised and revised items by incorporating them into a new LEAP O-3 instrument. To limit instrument length, three parallel forms of the LEAP O-3 should be developed, each adding five of the 15 items developed for each scale.

5. Conduct test-retest analyses of the LEAP O-3 to establish the stability of this measure using samples drawn from the 1993 ROTC applicant pool. Experience gained thus far on the O-2D (ROTC) have already reached acceptable levels (.73).

6. Refine the ALS Ordinal empirical key used in the LEAP O-2D. Its utility in improving the predictive efficiency of the LEAP over that obtained using the rational key was not consistently demonstrated. Empirically keyed LEAP scores correlated no better with peer rating dimensions than did rationally keyed LEAP scores. Test-retest results based on the empirical key yielded lower reliability coefficients than did results based on the rational key.

7. Devise appropriate algorithms to determine which LEAP component scales need correction for faking, and to what degree.

8. Adapt the LEAP for use by the other DOD departments, especially the Army, since the Army is the lead service in leadership research.

Steps for further validation of the LEAP instrument are:

1. Secure access from higher headquarters to administer LEAP O-3 (ROTC) to collegiate ROTC applicants.

2. Initiate a longitudinal study of ROTC selectees using the 1991 Summer Encampment Participants and gather criterion data as they mature.

3. Use the results of a LEAP O-3 (ROTC) field testing to make hypothetical "go - no go" selection decisions entirely independent of the operational selection process.

4. Compare the hypothetical selection decisions to actual performance for successful applicants as performance benchmarks (e.g., graduation/ nongraduation, Distinguished Graduate status) become available. The selection "hit rate" for selectees chosen without LEAP data should be compared with the "hit rate" achieved when LEAP results are included in the decision-making process and when used alone.

5. Field test and refine the LEAP O-2D (OTS) instrument using respondents from the OTS applicant pool. As with the ROTC population, LEAP data gathered would be used for research purposes only and not as part of the selection process.

6. Greatly expand the criteria used for validation.

7. Field test new versions of the LEAP with other DOD and civilian subjects.

REFERENCES

Appel, V.H., Grubb, P.D., Shermis, M.D., Watson, T.W., & Cole, R.W. (1990). *Initial development of a Biographical Survey for Use in Selection and Classification* (AFHRL-TR-90-19, AD-B146 272L). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Appel, V.H., Grubb, P.D., Elder, E.D., Leamon, R.E., Watson, T.W., & Earles, J.A. (1991). *Leadership effectiveness profile (LEAP): Organizational Taxonomy Enlisted Pool Development* (AL-TP-1991-0025, AD-A239 973). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Bass, B.M. (1985). *Leadership and performance beyond expectations*. New York: The Free Press.

Berger, F.R., Gupta, W.B., Berger, R.M., & Skinner, J. (1988). *Air Force Officer Qualifying Test (AFOQT) Form P: Test Construction.* (AFHRL-TR-88-30, AD-A200 678). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Berger, F.R., Gupta, W.B., Berger, R.M., & Skinner, J. (1990a). *Air Force Officer Qualifying Test (AFOQT) Form P: Test Manual.* (AFHRL-TR-89-56, AD-A221 004). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Berger, F.R., Gupta, W.B., Berger, R.M., & Skinner, J. (1990b). *Prototypes of cognitive measures for Air Force officers: Test development and item banking* (AFHRL-TP-89-73, AD-A222 670). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Bunch, M.B. (1974). *Retest reliability of the University of Georgia biographical questionnaire*. Unpublished master's thesis. University of Georgia, Athens, Georgia.

Carretta, T.R. (1987). *Basic Attributes Test (BAT) system: Development of an automated test battery for pilot selection* (AFHRL-TR-87-9, AD-A185 649). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Cascio, W.F., & Ramos, R.A. (1986). Development and application of a new method for assessing job performance in behavioral/economic terms. *Journal of Applied Psychology, 71*, 20-28.

Conger, J.A. (1989). *The charismatic leader.* San Francisco: Jossey-Bass, Inc.

Conger, J.A., & Kanungo, R.N. (1988). *Charismatic leadership*. San Francisco: Jessey-Bass, Inc.

Conger, J.A., & Kanungo, R.N. (Eds.) (1988). *Charismatic leadership*. San Francisco: Jossey-Bass, Inc.

Cronbach, L.J. (1970). *Essentials of psychological testing* (3rd Ed). New York: Harper and Row.

Doll, R.E. (1971). Item susceptibility to attempted faking as related to item characteristic and adopted faking set. *The Journal of Psychology, 77*, 9-16.

Driskell, J.E., & Olmstead, B. (1989). Psychology and the military: Research applications and trends. *American Psychologist, 44* (1), 43-54.

Elliott, L., & Watson, T.W. (1987). How useful can we expect assessment center derived measures of leadership and managerial ability to be? In V.H. Appel (Chair), *New directions in leadership and managerial selection*. Symposium presented at Texas Psychological Association.

Greenacre, M.J. (1984). *Theory and applications of correspondence analysis*. Orlando, FL: Academic Press.

Kyllonen, P.C. (in press). CAM: A theoretical framework for cognitive abilities measurement. In D. DeHerman (Ed.), *Current topics in human intelligence: Theories of intelligence*. Norwood, NJ: Ablex.

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher, 18*, 5-11.

Morales, M.M. (1991). *The function of general and specific abilities in the Air Force Officer Qualifying Test*. Unpublished master's thesis, St. Mary's University, San Antonio, TX.

Mumford, M.D., & Owens, W.A. (1987). Methodological review: Principles, procedures, and findings in the application of background data measures. *Applied psychological measurement, 11* (1), 1-31.

Owen, W.A. (1976). Background data. In M.D. Dunnette (Ed), *Handbook of industrial and organizational psychology* (pp. 609-644). New York: John Wiley.

Ree, M.J., & Earles, J.A. (1990a). *Differential validity of a differential aptitude* (AFHRL-TR-89-59, AD-A222 190). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Ree, M.J., & Earles, J.A. (1990b). *Estimating general ability in the ASVAB: The three faces of g* (AFHRL-TR-90-38, AD-227 710). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Ree, M.J., & Earles, J.A. (1990c). *The relationship among general ability, specific ability, and job category for predicting training performance* (AFHRL-TR-90-46, AD-A228 051). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Robertson, I.T., & Smith, M. (1989). *Advances in selection and assessment*. New York: John Wiley & Sons.

Saunders, V.C. (1983). *Verification of responses to an autobiographic data form*. Unpublished master's thesis, University of Georgia, Athens, GA.

Shaffer, G.S., Saunders, V., & Owens, W.A. (1986). Additional evidence for the accuracy of biographical data: Long term retest and observer ratings. *Personnel Psychology*, *39*, 791-809.

Siem, F.M. (1990). *Predictive validity of an automated personality inventory for Air Force pilot selection* (AFHRL-TP-90-55, AD-A226 731). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Skinner, J., & Ree, M.J. (1987). *Air Force Officer Qualifying Test (AFOQT): Item and Factor Analysis of Form O* (AFHRL-TR-86-68, AD-A184 975). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Sparks, C.P. (1988). *My love affair with biodata: A defense of empiricism*. Paper presented at the 96th Annual Convention of the American Psychological Association, Atlanta, GA.

Trent, T.T. (1987). *Armed Forces adaptability screening: The problem of item response distortion*. Paper presented at the 95th Annual Convention of the American Psychological Association, NY.

Watson, T.W. (1989). The Leadership Effectiveness Assessment Profile (LEAP): Prototype development and future plans. *Proceedings of the 31st Annual Conference of the Military Testing Association*, 397-402.

Watson, T.W., & Besetsny, L.K. (1991). Present and proposed nontraditional measures developed by the Air Force. *Proceedings of the 33rd Annual Conference of the Military Testing Association*, 571-576.

Watson, T.W., & Besetsny, L.K. (1992). Leadership research at the Air Force Human Resources Directorate. *Proceedings of the 34th Annual Conference of Military Testing Association*.

Watson, T.W., Elliott, L.R., & Appel, V.H. (1988, April). Enhancing the selection of Air Force personnel: New research directions. In *Proceedings, Psychology in the Department of Defense Eleventh Symposium* (USAFA-TR-88-1, pp. 120-124). Colorado Springs, CO: Department of Behavioral Sciences and Leadership, United States Air Force Academy.

Weller, S.C., & Romney, A.K. (1990). *Metric scaling*. Sage University Paper Series on Quantitative Applications in the Social Sciences No. 75. Beverly Hills, CA: Sage.

Welsh, J.R., Jr., Watson, T.W., & Ree, M.J. (1990). *Armed Services Vocational Aptitude Battery (ASVAB): Predicting military criteria from general and specific abilities* (AFHRL-TR-90-63, AD-A229 191). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Winter, D.G. (1987). Leader appeal, leader performance, and the motive profiles of leaders and followers: A study of American presidents and elections. *Journal of Personality and Social Psychology*, *52* (1), 196-202.

Young, F.W. (1981). Quantitative analysis of qualitative data. *Psychometrika*, *46*(4), 357-388.

APPENDIX A

DEMOGRAPHIC CHARACTERISTICS FOR LEAP O-2C AND LEAP O-2D

Demographic Characteristics for LEAP O-2C (OTS), (n = 138)

Variable	Frequency	Percentage
Age		
22 or under	2	1.4
23-24	33	23.9
25-26	38	27.5
27-28	20	14.5
29-30	27	19.6
Over 30	18	13.0
Gender		
Male	120	87.0
Female	18	13.0
Ethnicity		
American Indian	1	0.7
Asian	3	2.2
Black	2	1.5
Hispanic	1	0.7
White	130	94.9
Marital Status		
Married	72	52.2
Involved in an enduring relationship	0	0.0
Single	60	43.5
Other	6	4.3
Family Total Income (while in High School)		
Under $10,000	4	3.1
$10,001 - $15,000	10	7.9
$15,001 - $20,000	12	9.4
$20,001 - $30,000	22	17.3
$30,001 - $40,000	25	19.7
$40,001 - $50,000	19	15.0
$50,001 - $70,000	23	18.1
$70,001 - $100,000	10	7.9
Over $100,000	2	1.6
Do not know	0	0.0

Demographics Characteristics for LEAP O-2 (OTS): (Continued)

Variable	Frequency	Percentage
Year Graduated from High School		
1986	4	2.9
1985	16	11.6
1984	26	18.8
1983	19	13.8
1982 or earlier	73	52.9
Years of Prior Service		
None	61	44.2
1-2	2	1.4
3-5	12	8.7
6-8	31	22.5
9-10	13	9.4
Over 10	19	13.8
Size of Primary Community Lived In		
Fewer than 1,000 residents	6	4.3
1,001 to 5,000	11	8.0
5,001 to 10,000	11	8.0
10,001 to 25,000	17	12.3
25,001 to 50,000	26	18.8
50,001 to 100,000	18	13.0
100,001 to 200,000	14	10.1
200,001 to 500,000	0	0.0
500,001 to 1,000,000	0	0.0
More than 1,000,000	0	0.0
Father's Highest Education		
Fewer than 8 years	5	3.6
9-12 years, but not a H.S. graduate	4	2.9
High school graduate	37	26.8
1-2 years of college	18	13.0
3-4 years of college, but not coll. grad.	4	2.9
College graduate	42	30.4
Master's level degree	19	13.8
Doctoral level degree	9	6.5

Demographics Characteristics for LEAP O-2 (OTS): (Concluded)

Variable	Frequency	Percentage
Mother's Highest Education		
Fewer than 8 years	1	0.7
9-12 years, but not a H.S. graduate	7	5.1
High school graduate	52	37.7
1-2 years of college	30	21.7
3-4 years of college, but not coll. grad.	7	5.1
College graduate	32	23.2
Master's level degree	8	5.8
Doctoral level degree	1	0.7
Overall College GPA		
4.0 - 3.8	29	21.0
3.7 - 3.5	33	23.9
3.4 - 3.0	49	35.5
2.9 - 2.5	22	15.9
2.4 - 2.0	5	3.6
1.9 or less	0	0.0
GPA in College Major		
4.0 - 3.8	53	38.4
3.7 - 3.5	41	29.7
3.4 - 3.0	30	21.7
2.9 - 2.5	12	8.7
2.4 - 2.0	2	1.4
1.9 or less	0	0.0

Demographic Characteristics for LEAP O-2D (ROTC), (n = approximately 673)

Variable	Frequency	Percentage
Age		
17 and under	1	0.2
18	4	0.6
19	87	13.2
20	300	45.4
21	113	17.1
22	56	8.5
Over 22	100	15.1
Gender		
Male	547	81.2
Female	127	18.8
Ethnicity		
American Indian	4	0.6
Asian	21	3.1
Black	42	6.1
Hispanic	36	5.4
White	560	83.3
Other	10	1.5
Non-white	113	16.7
Marital Status		
Single	582	87.9
Living with a partner	15	2.2
Married	59	9.1
Separated/Divorced	3	.4
Divorced, now remarried	2	.3
Family Total Income (while in High School)		
Under $10,000	21	3.0
$10,001 - $15,000	21	3.1
$15,001 - $20,000	39	5.8
$20,001 - $30,000	76	11.3
$30,001 - $40,000	109	16.1
$40,001 - $50,000	122	18.2
$50,001 - $70,000	136	20.3
$70,001 - $100,000	63	9.2
Over $100,000	29	4.3
Do not know	58	8.6

Demographic Characteristics for LEAP O-2D (ROTC): (Continued)

Variable	Frequency	Percentage
Year Graduated from High School		
1991	1	0.1
1990	350	52.8
1989	135	20.7
1988	62	9.4
1987	113	17.0
1986 or earlier	0	0.0
Size of Primary Community Lived In		
Fewer than 1,000 residents	33	5.1
1,001 to 5,000	64	10.7
5,001 to 10,000	53	7.9
10,001 to 25,000	100	14.9
25,001 to 50,000	94	14.0
50,001 to 100,000	100	15.2
100,001 to 200,000	56	8.5
200,001 to 500,000	58	9.0
500,001 to 1,000,000	62	9.3
More than 1,000,000	35	5.4
Region Lived In		
Northeast	141	21.5
Southeast	139	21.0
North Central	119	18.2
South Central	80	12.3
Northwest	75	11.4
Southwest	101	15.5
Father's Highest Education		
Fewer than 8 years	12	1.8
9-12 years, but not a H.S. graduate	15	2.2
High school graduate	131	19.5
1-2 years of college	139	20.7
3-4 years of college, but not coll. grad.	37	5.5
College graduate	164	24.4
Master's level degree	136	20.2
Doctoral level degree	39	5.8

Demographic Characteristics for LEAP O-2D (ROTC): (Concluded)

Variable	Frequency	Percentage
Mother's Highest Education		
Fewer than 8 years	1	1.2
9-12 years, but not a H.S. graduate	24	3.6
High school graduate	169	25.7
1-2 years of college	162	24.6
3-4 years of college, but not coll. grad.	43	6.4
College graduate	169	25.9
Master's level degree	80	12.1
Doctoral level degree	3	0.4

APPENDIX B

LEAP SCALING PROJECT REPORT

LEAP Scaling Project Report

Introduction

This project grew out of an attempt to improve the psychometric properties of the LEAP biodata survey instrument. Biodata traditionally has produced predictive validity coefficients ranging in the high r=.20's for most industrial or business criterion measures and has generally been considered a good predictor when combined with other relevant variables. Robertson and Smith (1989), using meta-analytic techniques, synthesized the large amount of data available on the validity of commonly-used predictors and Table 1 shows how each is related.

Table 1. Range of mean validity coefficients for commonly-used predictors of work or business success. Source: Robertson, I.T., & Smith, M. (1989). Personnel selection methods. In M. Smith and I Robertson (eds.), *Advances in Selection and Assessment*. New York: John Wiley & Sons.

Predictor	Range of Mean Validity Coefficients
Work Sample	.38 to .54
Ability Composite (General Mental Ability plus Psychomotor Ability)	.53
Assessment Center	.41 to .43
Supervisor/Peer Evaluation	.43
General Mental Ability	.25 to .45
Biodata	.24 to .28
References	.17 to .26
Interviews	.14 to .23
Personality Assessment	.15
Self-Evaluation	.15
Interest Assessment	.10

One of the reasons that predictive validity coefficients rarely exceed r=.60 has to do with the relationship between the validity and reliability coefficient. This relationship is summed up in Formula 1:

$$r_{xy} \leq \sqrt{r_{11}} \qquad \text{Formula 1}$$

where r_{xy} represents the predictive validity coefficient and r_{tt} represents a reliability coefficient.

This formula indicates that the reliability coefficient places a bound on what the validity coefficient can be. For example, if the reliability coefficient is r=.64, then the highest the validity coefficient can be is r=.80. Other factors which attenuate the validity coefficient include restrictions of range on the criterion variable, homogeneity of the population for which the measures are being taken, violations of the homoscedasticity assumption (i.e., equal variability about the regression line), and violations of the linearity assumption (i.e., the relationship can be characterized by a straight line).

Estimation of the reliability of biodata survey forms can be particularly troublesome. Typically the test-retest coefficient is the most appropriate measure of reliability since the behaviors (or behavior intentions) inventoried should be stable over time. However if the survey is based upon some underlying theoretical framework (i.e., rationally keyed), then the items generated to measure the framework (or subconstructs within the framework) should be internally consistent. The problem is that the bandwidth for behaviors measured by most biodata surveys is so wide that they tend not to converge (Cronbach & Gleser, 1965). For example, when measuring the construct "leadership," activity levels inventoried in sports do not necessarily overlap with activity levels in academics.

One potential way to enhance the internal consistency of the biodata survey form is to optimize the scaling technique used to score the instrument. This may involve a partial or full-fledged abandonment of the rational key. If the departure from rational keying is partial, then the "correct" answer remains the same, but credit given to "wrong"

answers, if any, is empirically reweighted. If the departure from the rational key is full-fledged, then both the "correct" answers and weightings given to "wrong" answers are empirically determined. The dilemma here is that bolstering the internal consistency of the scale's subconstructs may undermine the theoretical framework from which the scale was developed. Moreover, if the rational key is entirely abandoned (e.g., no underlying construct is purportedly being measured), then one might as well use the stability coefficient as the best estimate of reliability.

Optimization techniques are not without disadvantages. Oftentimes the resulting values from such an analysis will bear no relationship to the initial rational key leaving researchers with an instrument having little face validity. Additionally, the resulting set of weights, because they are optimized to a particular sample, may result in initially high validity and reliability coefficients only to measurably diminish upon cross-validation. These criticisms have resulted in the development of scaling techniques capable of reflecting both the charisteristics of sample data while incorporating various theory-based constraints.

The purpose of this study was to find or develop a scaling approach for biodata which would improve the internal consistency of the scales measured by the instrument, but at the same time provide for a rational key or response patterns consistent with a rationally-developed key.

Major Scaling Approaches

Scaling models can be differentiated a variety of ways. Classification schemes include (1) what is being scaled (e.g., persons, stimuli, or both persons and stimuli), (2) the item trace lines of the scaling model [e.g., normally distributed (Thurstone), cumulative normally distributed (Likert), or step function (Guttman)], (3) types of data (e.g., preference, single stimulus, stimulus comparison, or similarities), and (4) scale dimensionality (e.g., unidimensional or multidimensional). In addition, the scaling models can be taxonomized according to the underlying assumptions of the measurement model. Measurement models are available for both classical test theory and item response theory (IRT).

Both classical and IRT models were investigated for the purposes of this study. IRT was abandoned early in the investigation process despite the fact that it conceivably would have some advantages over classical test theory. For example, data calibrated using IRT are thought to be "sample independent". That is, data calibrated from one sample of potential officer candidates would be generalizable to subsequent samples, assuming that the population of officer candidates remains relatively stable. However, most IRT models assume that the scale being calibrated is unidimensional, is composed of a relatively large number of items, and is calibrated on a large pool of subjects. While the LEAP is thought to be composed of relatively independent subscales, each subscale is measured by only 7-15 items, which is generally considered too few (Hambleton & Swaminathan, 1985). Moreover, while multidimensional IRT model have been proposed, there is no consensus among researchers as to how they might actually be implemented. Even the assumption of unidimensionality at the subscale

level is questionable, given the relatively low internal consistency values obtained through classical measurement theory. Finally, IRT assumes that an *a priori scoring key, rational or empirical, exits. The purpose of this study was,* of course, to generate an optimal scale. There is limited research on attempts to generate optimal scales based on IRT.

A number of classical scaling approaches were investigated as well. In addition to the traditional Thurstone, Likert, and Guttman scaling approaches, the following scaling techniques, among others, were studied for their application potential:

(1) Categorical Optimal Scaling--maximizes the cannonical correlation between the item and the criterion.

(2) Conjoint Analysis--provides weights for items and their alternatives from a utility perspective.

(3) Coombs Unfolding Technique--ranking of items to see if the resulting scale may be translated to a common judgment scale.

(4) Discriminant Analysis-reverse of the traditional correlational methods. One uses a criterion to predict to categories.

(5) Factor Analysis--use loadings on interpreted factors as weights for the dichotomized items.

The techniques associated with categorical optimal scaling (Goodman, 1984) appeared to hold the best promise for the purposes of this project. These approaches not only were geared to maximized the homogeneity of the resulting scale, but met the data assumptions as well. With respect to the data assumptions, a decision was made to treat the biodata as either nominal (the "correct" answer is empirically determined) or ordinal (the "correct" answer is rationally determined but the remaining response alternatives are empirically reweighted).

What follows is a description of two general classes of categorical optimal scaling: Alternating Least Squares Optimal Scaling (ALSOS) and Correspondence Analysis.

Alternating Least Squares Optimal Scaling

The phrase "optimal scaling" is a ubiquitous one which refers to the process by which one assigns numerical values to observation categories in a way which maximizes the relationship between the observations and the data analysis model at a given scale of measurement (Bock, 1960). While there are several algorithms which can be applied to find optimal solutions, the use of the Alternating Least Squares (ALS) has advantages in that they can describe qualitative data by quantitative models falling into three general classes: (a) The General Linear Model; (b) the Component (Factor) Model; and (c) The General Euclidean Model.

The ALS algorithms work by dividing all of the parameters into two mutually exclusive and exhaustive subsets: (a) the parameters of the model; and (b) the parameters of the data (i.e., optimal scaling parameters). The algorithms then optimize a loss function by alternately optimizing with respect to one subset, then the other. The optimization proceeds by obtaining the least squares estimates of the parameters in one subset while assuming that the parameters in all other subsets are constants. This is often referred to as a conditional least squares estimate, since the least squares nature is conditional on all the values of the parameters in the other subsets. Once a conditional least squares estimate has been obtained, the old estimates of the parameters are replaced by the new estimates. The algorithm then switches to another subset of parameters (i.e., each of the two subsets may itself contain parameter subsets) to obtain their conditional least

squares estimates. The iterations continue until convergence takes place. The only drawback is that the ALS procedure does not guarantee convergence on the globally least squares solution, but rather guarantees convergence on a particular type of local least squares solution. The local optimum is determined by the initilization process. As a way to address this problem most algorithms are initialized by applying a least squares procedure to the raw data under the assumption that the raw data are quantitative, as the user has coded them (Young, 1981).

In order to describe the process in more detail it is necessary to define a column vector of n raw observations. This observed vector is denoted as $\mathbf{o}$, with general element o_i. (Boldface lower case letters refer to column vectors, and italicized lower case letters to scalars). The model estimates $\hat{\mathbf{z}}$, are defined with general element $\hat{z}_i$, and the optimally scaled observations $\mathbf{z}^*$, with general element z_i^*. The elements of $\mathbf{o}$ are organized so that all observations in a particular category are contiguous. The elements of $\hat{\mathbf{z}}$ and $\mathbf{z}^*$ are organized in a fashion having a one-to-one correspondence with the elements of $\mathbf{o}$. The element z_i^* is the parameter representing the observation o_i. The vector $\hat{\mathbf{z}}$ is called the "model estimates" because it is the model's estimates, in the least squares sense, of the optimally scaled data $\mathbf{z}^*$.

The transformation $\mathcal{t}$ (script letters indicate transformations) of the raw observations which generates the optimally scaled observations,

$$\mathcal{t}[\mathbf{o}] = [\mathbf{z}^*] \qquad \text{Formula 2}$$

where the precise definition of $\mathcal{t}$ is a function of the measurement characteristics of the observations, and exists as a least squares relationship between the model's estimates of the scaled data ($\hat{\mathbf{z}}$) and the actual scaled data ($\mathbf{z}^*$), given that the measurement characteristics

of **o** are strictly maintained. The value assigned to z_i^* is the optimal parameter value for the observation o_i (Young, 1981, p. 362).

For the situation where the data are treated as nominal or ordinal, but an underlying distribution is continuous, the process restriction

$$t^{c}(o_i \sim o_m) \rightarrow (z_i^- = z_m^-) \leq \begin{Vmatrix} z_i^* \\ z_m^* \end{Vmatrix} \leq (z_i^+ = z_m^+) \qquad \text{Formula 3}$$

is applied where ~ indicates empirical equivalence (i.e., membership in the same category) and z_i^- and z_i^+ are the lower and upper bounds of the interval of real numbers. One of the implications of empirical (categorical) equivalence is that the upper and lower boundaries of all observations in a particular category are the same for all the observations (Young, 1981, p. 364).

ALSOS (Nominal). In order to estimate data parameters, it is necessary to introduce one final component referred to here as an *indicator matrix* with elements defined as follows:

$$u_{\text{pic}} = \begin{cases} 1 \text{ iff } o_i \in \text{category c} \\ 0 \text{ otherwise} \end{cases}$$

where U_p is defined as an (n x n_c) binary matrix with a row for each of the n observations in partition p, and column for each of the n_c categories. For convenience the subscript p will be left off when referring to the indicator matrix.

The minimization function for nominal-continuous data is

$$t^{m}_{\mathbf{Z^u}} = \mathbf{U}(\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}'\hat{\mathbf{z}} \qquad \text{Formula 4}$$

where $\mathbf{Z^u}$ represent the unnormalized scale data and the continuous process restriction $[t^n, (3)]$ (i.e., that each optimally scaled observation should reside in some interval) is imposed. This formula places no restrictions on the formation of the intervals. $\mathbf{U}'\mathbf{U}$ is a diagonal (n_c x n_c) matrix with a row and column for each observation category, $\mathbf{U}'\hat{\mathbf{z}}$ is an n_c element vector with the sum of the $\hat{z}_j$'s as its elements, and

$(U'U)^{-1}U'\hat{z}$ is an n_c element column vector with the mean of the appropriate $\hat{z}_j$'s as its elements (Young, 1981, p. 369).

The n_c data parameters for the partition under consideration are normalized using the following transformation:

$$z^* = \left[\frac{||\hat{z}||}{||z^u||}\right] \qquad \text{Formula 5}$$

ALSOS (Ordinal). If the data are treated ordinally, then a "correct response" can be designated and the remaining responses empirically reweighted. These variables have, in addition to the process restraints, the restriction that the real numbers assigned to observations in different categories represent the order of the empirical observations such that

$$t^o: (o_i \vee o_m) \rightarrow (z_i^* \leq z_m^*) \qquad \text{Formula 6}$$

where the superscript on t^o indicates the order restriction, and where $\vee$ indicates empirical order.

The minimization function for the continuous ordinal optimal scale is given as

$$t^{oo}: z^u = U(U'U)^{-1}P'\hat{z} \qquad \text{Formula 7}$$

where P is Kruskal's (1964) primary least squares monotonic transformation. The matrix P is a binary (n x n) block-diagonal permutation matrix. It has n_b blocks, each of which has an order equal to the corresponding element of $U'U$. Each block represents a permutation matrix having a single one in each row and column. P has only zeros outside the blocks. The matrix $U'U$ is interpreted as the number of observations is in each block and $(U'U)^{-1}P'\hat{z}$ contains the unnormalized least squares observation category parameter estimates.

For ordinal continuous data, the parameters are transformed in the same manner as for continuous nominal data. Missing data for both the

nominal and ordinal situation are coded in U as though they are each in a separate category (Young, 1981, p. 370).

Correspondence Analysis

Correspondence Analysis (Greenacre, 1984), primarily viewed as a graphical approach to the analysis of nominal data, provides optimal weights based on the dimensionality of x and y. This approach, based on the fundamental singular values decomposition of a matrix, has been discovered and rediscovered across multiple disciplines since the early efforts of Fisher (1940) and Guttman (1941) who were apparently unaware of each other's work. The technique has alternatively been referred to as optimal scaling, dual scaling, Guttman Scaling, Pattern Analysis, etc. (cf. Weller & Romney, 1990). Regardless of name, the technique resembles factor analysis though assumes a nominal scale of measurement allowing orthogonal dimensions among variables to be identified and rescaled. For each dimension among the variables, a set of optimal weights are generated allowing the researcher to isolate relationships among variables not available through other optimization techniques. Because of its inherent multidimensional nature, biodata may benefit especially from the use of Correspondence Analysis.

The first step in correspondence analysis is to normalize the data by dividing each row entry by the square root of the product of corresponding row and column totals. Notationally this is written

$$h_{ij} = \frac{f_{ij}}{\sqrt{f_{i.} f_{.j}}} \qquad \text{Formula 8}$$

where h_{ij} is the entry for a given cell, f_{ij} is the original cell frequency, $f_{i.}$ is the total count for row i, and $f_{.j}$ is the total count in column j. In matrix notation, this can be expressed as

$$H = S^{-1/2} F C^{-1/2} \qquad \text{Formula 9}$$

where **H** contains the transformed matrix, **F** is the frequency matrix, and $S^{-1/2}$ and $C^{-1/2}$ are diagonal matrices whose entries consist of reciprocals of the square root of the row marginal totals and column marginal totals, respectively (Weller & Romney, 1990, p. 60).

In the second step, the basic structure of the normalized **H** matrix is found using the singular value decomposition (SVD) technique. Singular value decomposition of a matrix is a common mathematical approach to reducing a matrix to its elemental row and column components.

Any matrix A can be decomposed into the basic structure:

$$A_{mxn} = (P_{mxk})(\Delta_{kxk})(Q'_{kxn}) \qquad \text{Formula 10}$$

where $k \leq$ the minimum of rows or columns (m or n), A is the result matrix, P and Q are orthonomals by columns, and Delta is a diagonal with ordered positive entries. These values rescaled are the optimal scores for the resulting dimensions (where the number of dimensions is equal to the lesser of rows or columns minus 1). These optimal scores have the property of maximizing the canonical correlation between the two variables that are not viewed as having any particular scale of measurement. The first singular value is always one and successive values constitute canonical correlations or singular values.

The last step is to rescale the row (**U**) and column (**V**) vectors to obtain the canonical or optimal scores using the following formulas:

$$X_i = U_i\sqrt{f_{..}f_{i.}} \qquad \text{Formula 11}$$

$$Y_j = V_j\sqrt{f_{..}f_{.j}} \qquad \text{Formula 12}$$

where **X** and **Y** respectively represent the row and column vectors of canonical or optimal scores. It should be noted that the first vector of scores are all 1.0 and correspond to the independence model of Chi-square expected values. This vector is ignored an any subsequent analysis (Weller & Romney, 1990, pps. 60-61).

Monte Carlo Simulations

The simulation of empirical keying techniques sought to determine the relative merits of the three keying approaches under the three conditions discussed below. Each of these conditions was simulated with a wide range of interdependence between rows (criterion) and columns (items) for a total of 200 samples each (n=200). For practical purposes this may be thought of as multiple samples expressed in contingency table form, some having total independence to others having total dependence between rows and columns. Dependence between rows and columns may be thought of as reflecting an item's ability to differentially attract respondents: when no dependence exists, items are generally poor. High dependence indicates that particular alternatives will 'pull' homogenous groups of respondents thereby allowing the establishment of a distinct empirical weight.

In condition one, rational weights were defined to be equal with population alternative weights. In this instance it was hypothesized that the ordinal approach would perform as well as both nominal and Correspondence Analysis approaches. For each set of 200 samples generated in this simulation, sample frequency data were generated in the form of a 4 x 4 (item by criterion) frequency matrix that served as a contingency table having a varying degree of independence between rows and columns. In this way each keying approach could be evaluated against data having anywhere from 1) no relation between rows and columns (equal probability of assignment to any of the 16 cells), or 2) high relation between rows and columns. In condition 1 where rational weights were defined to be equal with population alternative weights high dependence between rows and columns was defined as placing a higher probability of frequency assignment along the primary diagonal.

In condition two, rational weights were defined to be the opposite of the population alternative weights. In this case high dependence between rows and columns was accomplished through assigning a higher probability of data to occur in the secondary diagonal (4,1; 3,2; 2,3; 1,4) indicating that initial weights were opposite of actual population parameters. Here it was hypothesized that nominal and CA approaches might do equally well but that the ordinal approach would result in poor estimation in tables where rows and columns were strongly dependent. Again, a range of contingency tables were generated from no dependency, to high row by column dependency.

In condition three, the usual range of row by column dependencies were generated with high row by column dependencies characterized by high probability of assignment to each of the 4 corners in the 4 by 4 table. In this way a multidimensional relationship was generated in the population which is typically the case with biodata type items. It was hypothesized again that the ordinal approach would fare poorly, with the nominal and approach achieving less satisfactory results than the Correspondence Analysis approach that was able to provide 2 sets of optimal weights consistent with the multidimensionality of the data.

Analysis Software

The simulations were written using the SAS IML (Interactive Matrix Language) product in conjunction with the SAS BASE product that allowed utilization of macro processing routines. Briefly, data were generated from a random normal deviate allowing nonrandom cell assignment according to probabilities from a normal distribution. Data generated were subjected to three scaling approaches resulting in a total of four scale values to examine. The four sets of scaled data were then correlated with the criterion with these results output for summary and

evaluation. The basic pattern of raw data generation, scaling and analysis was placed within loops to replicate the process a multiple number of times under differing conditions.

Results

As can be seen from the figures, the stated hypothesis were generally supported. In each of Figures 1 through 3 independence was defined as the chi-square statistic associated with the test of independence between rows and columns. These chi-square values constitute the x-axis while the y-axis contains the resulting correlation between the optimally weighted item and the criterion. In evaluating a multiple line plot of this nature, one would expect that correlations would be low for low levels of dependence (low chi-square values) and high for high levels of dependence (high chi-square values). For each of the weighting approaches a line of fit was generated between the level of independence (x-axis) and the resulting correlation with the criterion.

Recall that in Condition 1, rational weights were defined to be equal with population alternative weights. As can be seen from Figure 1 representing Condition 1, the nominal, ordinal and CA-1 approaches are roughly equal in performance across all levels of independence. The CA-2 algorithm, which incorporated two orthogonal dimensions rather than one, demonstrated low correlations with the criterion across all levels of independence. As expected, at high levels of independence, low correlations were obtained, while at high levels of dependence higher correlations were obtained. Note the superior performance of CA-1 at the high independence portion of the x-axis.

Figure 1. Correlation of the optimally weighted item and the criterion under different levels of dependence across nominal, ordinal, correspondence analysis-dimension 1, correspondence analysis-dimension 2 algorithms for Condition 1.

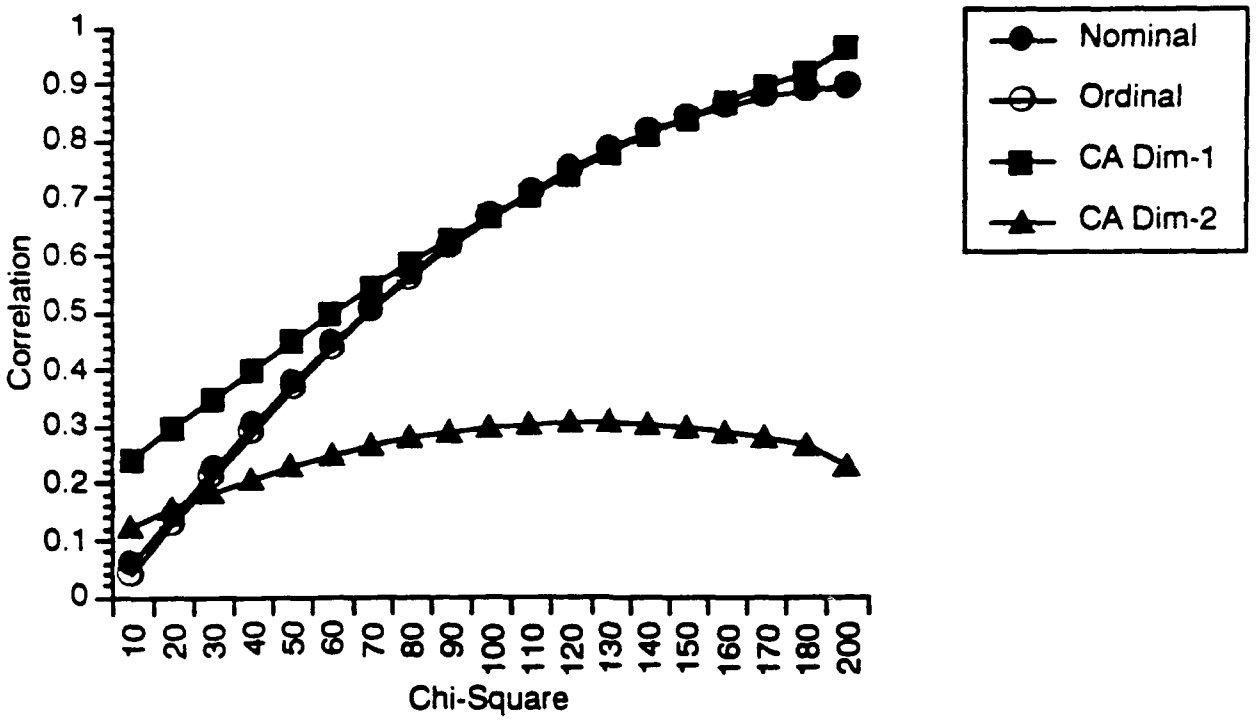
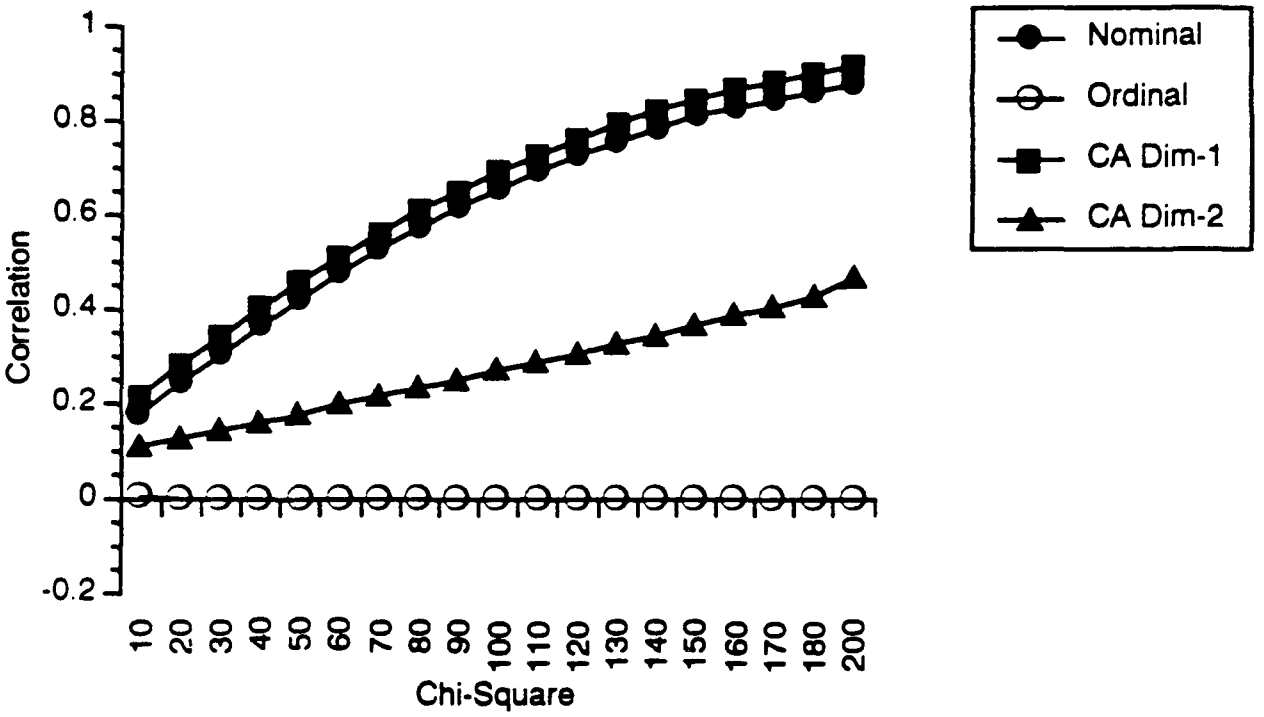


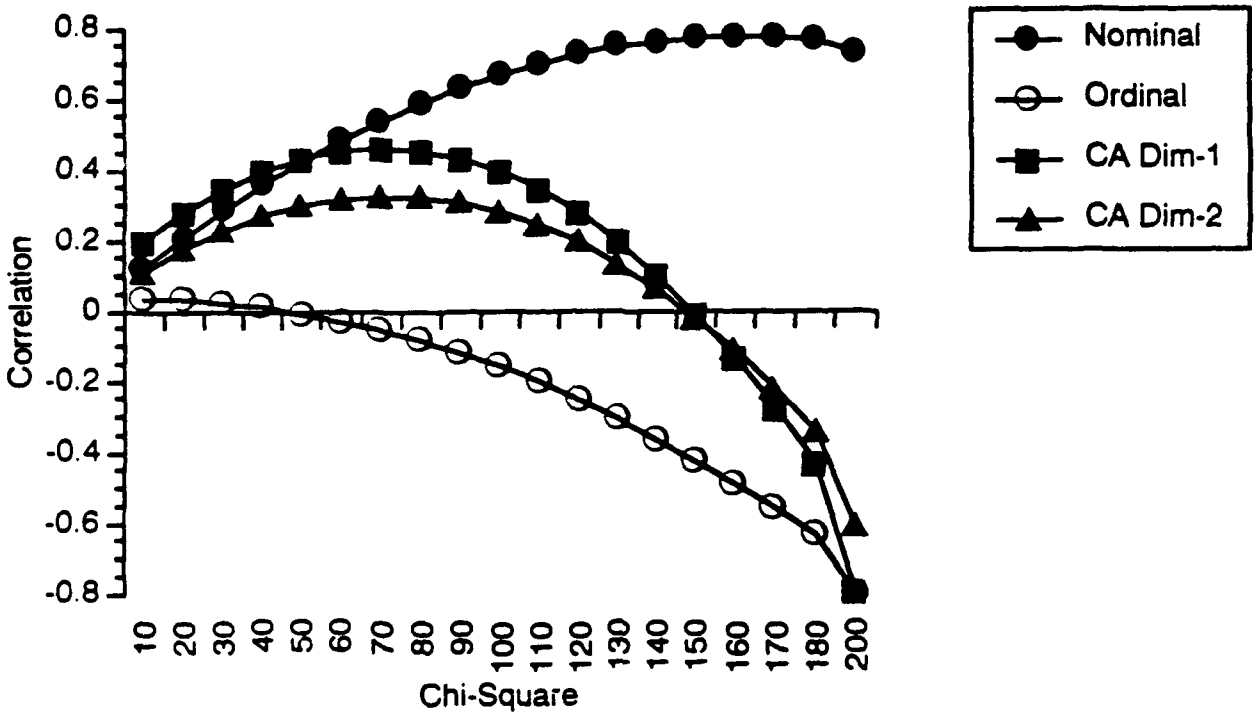
Figure 2 represents the correlations from the Condition 2 simulation. In this condition, rational weights were defined to be opposite of the population alternative weights. These regression lines support the hypothesis that the ordinal approach falls far short of both the nominal and CA approaches. That is, constraining weights based on a theory incompatible with actual population parameters results in weights having poor predictive utility. Note the differential performance of the nominal and ordinal approaches across all levels of independence.

Figure 2. Correlation of the optimally weighted item and the criterion under different levels of independence across nominal, ordinal, correspondence analysis-dimension 1, correspondence analysis-dimension 2 algorithms for Condition 2.



Finally, Figure 3 contains lines of fit for the results of the Condition 3 simulation. In this condition, the usual range of row by column dependencies were generated with high row by column dependencies characterized by high probability of assignment to each of the four corners in the four by four table. Here it is evident that only the nominal approach continues to provide a high degree of relationship between the optimally weighted item and the criterion variable. The ordinal, CA-1, and CA-2 conditions perform poorly and even demonstrate negative relationship in high dependence conditions.

Figure 3. Correlation of the optimally weighted item and the criterion under different levels of dependence across nominal, ordinal, correspondence analysis-dimension 1, correspondence analysis-dimension 2 algorithms for Condition 3.



Conclusions

While the expected results were confirmed for unidimensional situations in which there is a linear relationship between initial weights and population parameters, the picture is more complex for a multidimensional situation where no such relationship exists in the population.

The unidimensional situation was examined under two conditions. First when there is a positive correlation between the theoretical model and the weights given to the response alternatives, the nominal, ordinal, and CA-1 conditions performed equally well. With the CA-2 algorithm, the relationship was non-linear and became negative under high dependence situations. With the exception of the CA-2 condition, these results were predicted.

In the second unidimensional condition, the weights given the response alternatives were inversely related to the actual population parameters. It was predicted that the nominal and two correspondence analysis conditions would perform equally well. It was further predicted that the ordinal condition which is linked to the theoretical model would fare poorly in this condition. The results showed that the nominal and CA-1 algorithms did well, followed by the CA-2 algorithm. The ordinal algorithm, as predicted, had a near zero relationship between the optimal weight and criterion across the different levels of dependence.

In the third condition, no linear relationship existed between the population parameters and assigned optimal weights. This condition was designed to simulate one possible distribution of multidimensional data. In this situation, only the nominal data fared well across all levels of dependence. The other algorithms had fair performance at low levels of dependence, but did poorly at high levels of dependence.

Which algorithm one chooses to represent the data is primarily a function of how the data are distributed. If the data are unidimensional and weights correspond well to some underlying theory guiding item construction, then the nominal, ordinal, and CA-1 algorithms seem to be logical choices. The ordinal approach would be an attractive alternative in this situation in that the "correct" response would be based on theory while the remaining alternatives could be empirically weighted.

If the data are unidimensional, but there is no or an inverse correspondence of the optimal weights with the population weights, then the nominal or CA-1 algorithms appear to work well. This situation

corresponds to one in which no theory underlies item construction or that the theory guiding item writing is wrong.

In the situation where the data are multidimensional, only the nominal algorithm worked well. The two correspondence analysis algorithms performed acceptably well under low dependence, but broke down under high dependence. As expected, the ordinal condition did poorly under both low dependence and high dependence conditions.

Based on these simulations, it appears as if the nominal algorithm was the most robust with respect to maintaining the relationship between the optimal and population weights. While this was expected for unidimensional data, it was surprising how well the relationships were maintained under the multidimensional condition. It should be noted that this robustness might not be evident using other multidimensional distributions. The CA-1 algorithm worked well in the unidimensional conditions, but fared poorly with multidimensional data. The ordinal approach did well when the optimal weights corresponded with the population weights, but when this was not the case, or the data were multidimensional, the ordinal algorithm did not appear to be a good optimal scaling strategy. Surprisingly, the CA-2 algorithm did poorly across all three conditions. It was expected that adding the second orthogonal dimension might reduce its performance with unidimensional data, but pick up the multidimensionality of Condition 3. This did not happen. This may have occurred in part because Condition 3 represented just one of many possible multidimensional distributions.

The application of the simulations to LEAP data might be made in the following way. If one assumes that the scales are unidimensional (e.g., transformational leadership represents one hypothetical construct) and that the theory underlying item writing is a reasonably

good one, then the ordinal approach would be the most appealing optimal scaling algorithm. The weighting of the "correct" response would be consistent with the theory underlying item construction. The answers would be aligned with the rational key. If LEAP data are unidimensional, but the theory underlying item writing is faulty, then either the nominal or CA-1 algorithms would be most optimal scaling strategies.

There is some evidence that the LEAP scales are multidimensional. The internal consistency coefficients are modest indicating either a bandwidth fidelity problem or that the data have a multidimensional distribution. If the scales are in fact multidimensional, then the most optimal scaling algorithm would be one based on the nominal approach.

As can be seen from the above discussion, whether the data are unidimensional or multidimensional is critical to determining the preferred empirical scaling technique. Yet at the same time determining dimensionality is contingent on scoring the instrument, which, in turn, requires the selection of an empirical algorithm.

Caught between this chicken-and-the-egg dilemma, it was decided to make an *a priori* judgment about dimensionality based upon scale content. Using that perspective, undimensionality seemed most likely. Thus, for this project, the assumption has been made that the LEAP scales are unidimensional and that the theory underlying the items is sound. These assumptions support the use of the ordinal empirical key. Should these assumptions subsequently be found erroneous, an appropriate, alternate (e.g., nominal) scaling algorithm will need to be selected and applied.

References

Andrich, D. (1978). Relationships between the Thurstone and Rasch approaches to item scaling. *Applied Psychological Measurement, 3,* 446-460.

Andrich, D. (1978). Scaling attitude items constructed and scored in the Likert tradition. *Educational and Psychological Measurement, 38,* 665-680.

Alwin, D. F. (1973). The use of factor analysis in the construction of linear composites in social research. *Sociological Methods and Research, 2,* 191-214.

Asher, J. J. (1972). The biographical item: Can it be improved? *Personnel Psychology, 25,* 251-269.

Berkeley, M.H. (1953). A comparison between the empirical andrational approaches for keying a heterogeneous test. *USAF Human Resources Research Center Bulletin,* 53-54.

Bock, R. D. (1960). Methods and applications of optimal scaling. Psychometric Laboratory Report #25. Chapel Hill, NC: University of North Carolina.

Carroll, J.D., Green, P.E., & Schaffer, C.M. (1986). Interpoint distance comparisons in correspondence analysis. *Journal of Marketing Research, 23,* 231-280.

Cliff, N. (1973). *Scaling. Annual Review of Psychology, 24,* 473-506.

Clogg, C.C. (1977). *Unrestricted and restricted maximum likelihood latent structure analysis: A manual for users.* University Park: The Penn. State University, Population Issues Research Office.

Clogg, C.C. & Goodman, L.A. (1986). On scaling models applied to data from several groups, *Psychometrika, 51,* 123-135.

Clogg, C.C. & Goodman, L.A. (1984). Latent structure analysis of a set of multidimensional contingency tables. *JASA, 79,* 762-771.

Coombs, C.H. (1950). Psychological scaling without a unit of measurement. *Psychological Review, 57,* 145-158.

Cronbach, L.J., & Gleser, G.C. (1965). *Psychological tests in personnel decisions* (2nd Edition). Urbana: University of Illinois Press.

Dubois , P.H., Loevinger, J., & Gleser, G.C. (1952). The construction of keys for a biographical inventory (*Bulletin O, 52-18*). Lackland AFB, TX: USAF Human Resources Research Center.

England, G.W. (1961). *Development and use of weighted application blanks.* Dubuque, IA: William C. Brown.

Festinger, L. (1947). The treatment of qualitative data by scale analysis. *Psychological Bulletin, 44,* 149-161.

Goodman, L.A.. (1984). *The analysis of categorical data having ordered attributes.* New York: Harper and Row.

Gorden, R.L. (1977) *Unidimensional scaling of social variables: Concepts and procedures.* NY: Free Press.

Greenacre, M.J.. (1984). *Theory and applications of correspondence analysis.* Orlando, FL: Academic Press.

Guttman, L.L. (1947). A basis for scaling qualitative data. *American Sociological Review, 9,* 139-150.

Guttman, L.L.(1947). On Festinger's evaluation of scale analysis. *Psychological Bulletin, 44,* 451-465.

Hulin, C.L., Drasgow, F. & Parsons, C.K. (1983). *Item response theory.* Homewood, IL: Dow Jones-Irwin.

Klein, S.P., & Owens, W.A. (1965). Faking of a scored life history as a function of criterion objectivity. *Journal of Applied Psychology, 49,* 451-454.

Koch, W. R. (1983). Likert scaling using the graded response latent trait model, *Applied Psychological Measurement, 7,* 15-31.

Kruskal, J. B. (1964). Nonmetric multidimensional scaling. *Psychometrika, 29,* 1-27, 115-129.

Labovitz, S. (1970). The assignment of numbers to rank order categories. *American Sociological Review, 35,* 515-524.

Laurent, H. (1970). Cross-cultural cross-validation of empirically validated tests. *Journal of Applied Psychology, 54,* 417-423.

Lazarsfeld, P.F. & Henry, N.W. (1968). *Latent structure analysis.* Boston: Houghton-Mifflin.

Lecznar, W.B., & Dailey, J.T. (1950). Keying biographical inventories in classification test batteries. *American Psychologist, 5,* 279.

Likert, R. (1974). A method of constructing an attitude scale, (pp. 233-243) in G.M. Maranell (eds) *Scaling: A sourcebook for behavioral scientists.* Chicago: Aldine.

Likert, R. (1931). A technique for the measurement of attitudes. *Archives of Psychology.* NY: Columbia University Press.

Malloy, J. (1955). The prediction of college achievement with the life experience inventory. *Educational and Psychological Measurement, 15,* 170-180.

Malone, M.P. (1978). *Predictive efficiency and discriminatory impact of verifiable biographical data as a function of data analysis procedure.* Unpublished doctoral dissertation, University of Minnesota, Minneapolis.

Matteson, M. T., Osbur, H. G., & Sparks, C.P. (1969). A computer-based methodology for constructing homogeneous keys with applications to biographical data (*Personnel Psychology Services Center Report No. 1*). Houston, TX: University of Houston.

McIver, J.P. & Carmines, E.G. (1981). *Unidimensional scaling*. Newbury Park, CA: Sage.

Mitchell, T. W., & Klimoski, R. J. (1982). Is it rational to be empirical? A test of methods for scoring biographical data. *Journal of Applied Psychology, 67*, 411-418.

Mokken, R.J. (1971). *A theory and procedure of scale analysis*. The Hague: Mouton.

Mumford, M.D. & Owens, W.A. (1987). Methodology review: Principles, procedures, and findings in the application of background data measures, *Applied Psychological Measurement, 11*, 1-31.

Owens, W. A. (1976). Background biodata. In M. D. Dunnette (Ed.) , *Handbook of industrial and organizational psychology* (pp.610-644). Chicago, IL: Rand-McNally.

Perline, R., Wright, B.D., & Wainer, H. (1979). The Rasch model as additive conjoint measurement. *Applied Psychological Measurement, 3*, 237-256.

Schrader, A. D. (1975). *A comparison of the relative utility of several rational and empirical strategies for forming biodata dimensions.* Unpublished doctoral dissertation, University of Houston, TX.

Tenenhaus, M., & Young, F.W. (1985). An analysis ond synthesis of multiple correspondence analysis, optimal scaling, dual scaling, homogeneity analysis and other methods for quantifying categorical multivariate data, *Psychometrika, 50*, 91-119.

Thayer, P. W. (1977). Somethings old, somethings new. *Personnel Psychology, 30*, 513-524.

Torgerson, W. S. (1968). *Scaling International Encyclopedia of the Social Sciences, Vol. 14.* (pps.25-38). NY: Macmillan.

Weller, S. C., & Romney, A. K. (1990). *Metric scaling*. Sage University Paper Series on Quantative Applications in the Social Sciences No. 75. Beverly Hills, CA: Sage.

Wright, B.D. & Mead, R.J. (1976). BICAL: calibrating items with the Rasch model. *Research Memorandum No. 23*. Chicago: University of Chicago, Department of Education, Statistical Laboratory.

Young, F. W. (1981). Quantitative analysis of qualitative data. *Psychometrica, 46*(4), 357-388.

```
%macroits;
%let totn=40 ;     * total number of subs per xy ;
%let totitems=100; * total number  of items per gvar ;

*defines the number of items in each simulation
below the second number is the number of gvar levels ;
%let iv=%eval(&totitems*3);

proc iml ;
* counter for all items (regardless of var) ;
allitem=0;
*the row number in the matrices below is levels of var * totitems ;
*vectors to hold correlations ;
yraw=j(&iv,1,.);
ynom=j(&iv,1,.);
yord=j(&iv,1,.);
yca1=j(&iv,1,.);
yca2=j(&iv,1,.);
vchi=j(&iv,1,.);
gvar=j(&iv,1,.);

do vara= 2 to 8  by 3 ;
%do item=1 %to &totitems ;

yx=j(&totn,2,.);
allitem=allitem + 1 ;
* print allitem ;
x=. ; y=.;
do n=1 to &totn ;
   yran=int(rannor(0) * vara  + 8.5);
/* *==========block for perfect metric========;
   if yran =  1 then do; x=4; y=1; end; else
   if yran =  2 then do; x=2; y=4; end; else
   if yran =  3 then do; x=1; y=3; end; else
   if yran =  4 then do; x=3; y=1; end; else
   if yran =  5 then do; x=2; y=1; end; else
   if yran =  6 then do; x=2; y=1; end; else
   if yran =  7 then do; x=3; y=4; end; else
   if yran =  8 then do; x=1; y=1; end; else
   if yran =  9 then do; x=4; y=4; end; else
   if yran = 10 then do; x=3; y=3; end; else
   if yran = 11 then do; x=2; y=3; end; else
   if yran = 12 then do; x=4; y=3; end; else
   if yran = 13 then do; x=1; y=2; end; else
   if yran = 14 then do; x=3; y=2; end; else
   if yran = 15 then do; x=4; y=2; end; else
   if yran = 16 then do; x=1; y=4; end; else yran=99   ;
```

```
   *==========block for negative metric=======;
   if yran =  1 then do; x=4; y=4; end; else
   if yran =  2 then do; x=4; y=3; end; else
   if yran =  3 then do; x=3; y=4; end; else
   if yran =  4 then do; x=3; y=3; end; else
   if yran =  5 then do; x=2; y=4; end; else
   if yran =  6 then do; x=1; y=3; end; else
   if yran =  7 then do; x=2; y=3; end; else
   if yran =  8 then do; x=4; y=1; end; else
   if yran =  9 then do; x=1; y=4; end; else
   if yran = 10 then do; x=3; y=2; end; else
   if yran = 11 then do; x=3; y=1; end; else
   if yran = 12 then do; x=4; y=2; end; else
   if yran = 13 then do; x=2; y=2; end; else
   if yran = 14 then do; x=1; y=1; end; else
   if yran = 15 then do; x=1; y=2; end; else
   if yran = 16 then do; x=2; y=1; end; else yran=99   ;
     *==========block for corners ==============;
   if yran =  1 then do; x=2; y=3; end; else
   if yran =  2 then do; x=2; y=2; end; else
   if yran =  3 then do; x=3; y=4; end; else
   if yran =  4 then do; x=4; y=3; end; else
   if yran =  5 then do; x=2; y=4; end; else
   if yran =  6 then do; x=1; y=3; end; else
   if yran =  7 then do; x=1; y=1; end; else
   if yran =  8 then do; x=1; y=4; end; else
   if yran =  9 then do; x=4; y=1; end; else
   if yran = 10 then do; x=4; y=4; end; else
   if yran = 11 then do; x=3; y=1; end; else
   if yran = 12 then do; x=4; y=2; end; else
   if yran = 13 then do; x=2; y=1; end; else
   if yran = 14 then do; x=1; y=2; end; else
   if yran = 15 then do; x=3; y=3; end; else
   if yran = 16 then do; x=3; y=2; end; else yran=99   ;
   *==========end of metric types==============;
   yx[n,]=y||x ;
   if yran= 99 then n = n - 1 ;
   *print n y x yran ;
end;
*====make young & CA matrices to be filled==== ;
pyn=j(&totn,1,.);
pyo=j(&totn,1,.);

*=========the vectors for final CA analysis ==========;
* two cols for two optimal criterion dimensions ;
pcac=j(&totn,2,.);
pcac[,1]=yx[,1];
* two cols for two optimal dimensions ;
pcar=j(&totn,2,.);
pcar[,1]=yx[,2];
```

```
*========young analysis==========;
*nominal scaled with raw predictors ;
pyn=opscal(1,yx[,1],yx[,2]);
*ordinal to be scaled with ordered predictors;
pyo=opscal(2,yx[,1],yx[,2]);
* for constant result if min(pyo)=max(pyo) then pyo[&totn]=pyo[&totn-
1]+.11;
*====================end young analysis==================;

*====================ca matrix creation=================== ;
*1 is the minimum value in each cell ;
raw=j(4,4,1);
*fill contingency table with raw frequencies ;
do cb=1 to &totn;
  raw[yx[cb,2],yx[cb,1]] =
        raw[yx[cb,2],yx[cb,1]]+ 1 ; end;

*========generate chisq stat=============;
expec=j(4,4,.);  totraw=sum(raw);
diff =j(4,4,.);
rm = raw[,+]; cm=raw[+,];
do crm=1 to 4 ;
  do ccm=1 to 4 ;
    expec[crm,ccm]= (rm[crm]*cm[ccm])/totraw;
    diff[crm,ccm]=((raw[crm,ccm]-expec[crm,ccm])**2)/expec[crm,ccm];
  end; end;
chisq=sum(diff);
* print yx ;
dyx=yx-repeat(yx[:,],nrow(yx),1);
icorr=(dyx[,1]` * dyx[,2]) / sqrt(ssq(dyx[,1]) * ssq(dyx[,2]));
* print raw chisq icorr ;
*=======================================;

*====================ca optimal analysis============ ;

*========compute matrices to be filled==========;
norm=j(4,4,.);

*tables to hold final column and row scaled values;
pcca=j(4,3,.);
prca=j(4,3,.);
*=======compute marginal vectors and normalize raw data====== ;
rowm=raw[,+]; colm=raw[+,];
do rc=1 to 4 ;
  do cc=1 to 4;  *normalize the raw score matrix ;
    norm[rc,cc]=raw[rc,cc] / (sqrt(rowm[rc]*colm[cc])); end; end;

call svd(sumr,dsv,sumc,norm);
*RESCALE (normalize) sumr (row) components ;
do rc=1 to 4;
  do rcb=1 to 3 ;  *rescale the row components;
    prca[rc,rcb]=sumr[rc,rcb] * (sqrt(&totn/rowm[rc])); end; end;
*extract first and second predictor dimensions ;
prca=prca[,2]||prca[,3];
```

```
*RESCALE the col components;
do cc=1 to 4 ;
  do ccb=1 to 3 ;
    pcca[cc,ccb]=sumc[cc,ccb] * (sqrt(&totn/colm[cc])); end; end;
*extract first dimension criterion dimensions ;
pcca=pcca[,2]||pcca[,3];

*FILL raw y values with 2 sets of Y optimal values;
c=0;
do cfill=1 to 4; c = c + 1 ;
  do crows=1 to &totn;
     if pcac[crows,1]=cfill
        then pcac[crows,]=pcca[c,]; end; end;

*FILL raw x values with optimal values;
r=0;
do rfill=1 to 4 ; r= r + 1 ;
  do rose=1 to &totn;
     if pcar[rose,1]=rfill
         then pcar[rose,]=prca[r,]; end; end;
*=========end of CA==========;
* print yx pyn pyo pcac pcar ;

*=========generate corr matrix==========;
if max(pyn)=min(pyn) then pyn[nrow(yx)]=pyn[nrow(yx)]+.001 ;
if max(pyo)=min(pyo) then pyo[nrow(yx)]=pyo[nrow(yx)]+.001 ;
if max(pcac[,1])=min(pcac[,1]) then
pcac[nrow(yx),1]=pcac[nrow(yx),1]+.001 ;
if max(pcac[,2])=min(pcac[,2]) then
pcac[nrow(yx),2]=pcac[nrow(yx),2]+.001 ;
if max(pcar[,1])=min(pcar[,1]) then
pcar[nrow(yx),1]=pcar[nrow(yx),1]+.001 ;
if max(pcar[,2])=min(pcar[,2]) then
pcar[nrow(yx),2]=pcar[nrow(yx),2]+.001 ;
matx=yx||pyn||pyo||pcac||pcar ;
sum=matx[+,]; xpx=t(matx) * matx - t(sum) * sum/nrow(yx);
s=diag(1/sqrt(vecdiag(xpx))) ;
corr=s*xpx*s ;

yraw[allitem]=corr[1,2];
ynom[allitem]=corr[1,3];
yord[allitem]=corr[1,4];
yca1[allitem]=corr[7,5];
yca2[allitem]=corr[8,6];
vchi[allitem]=chisq;
gvar[allitem]=vara ;
* print corr ;

%end;  * item loop ;
end; * end of the var loop ;

*XXXXXXXXXXXXXXXXXXend of data processingXXXXXXXXXXXXXXXXXXXX ;

* print yraw ynom yord yca1 yca2 gvar vchi;
*==========jutput to DS==============;
create imlout var {     yraw ynom yord yca1 yca2 gvar vchi};
append var _all_ ;
```

```
proc sort   data=imlout ; by gvar ;

proc means data=imlout mean std n ;
        var yraw ynom yord yca1 yca2 ; by gvar ;
title "Corrs of Weights (Rational r=+1), Pop r=corners, N=&totn";
*/

goptions device=xbw;
symbol1 i=rcclm w=1;
symbol2 i=rcclm w=2;
symbol3 i=rcclm w=3;
symbol4 i=rcclm w=4;
symbol5 i=rcclm w=5;

proc gplot data=imlout ;
plot yraw*vchi  ;
title "Rational keyed (r=+1), Pop r=  1, N(item)=&totn";

proc gplot data=imlout ;
plot ynom*vchi  ;
title "Young-nominal method, Pop r=  1, N(item)=&totn";

proc gplot data=imlout ;
plot yord*vchi  ;
title "Young - ordinal method, Pop r=  1, N(item)=&totn";

proc gplot data=imlout ;
plot yca1*vchi  ;
title "CA dimension 1 method, Pop r=  1, N(item)=&totn";

proc gplot data=imlout ;
plot yca2*vchi ;
title "CA dimension 2 method, Pop r=  1, N(item)=&totn";

/*
goptions device=xbw;
Proc sort data=imlout ; by vchi ;
symbol1 i=rc   v=none w=1;
symbol2 i=rc   v=none w=2;
symbol3 i=rc   v=none w=3;
symbol4 i=rc   v=none w=4;
symbol5 i=rc   v=none w=5;
proc gplot data=imlout ;
plot yraw*vchi ynom*vchi yord*vchi yca1*vchi yca2*vchi / overlay ;
title "All methods (r=+1), Pop r=  1, N(item)=&totn";

proc reg data=imlout ;model yraw=vchi ;
proc reg data=imlout ;model ynom=vchi ;
proc reg data=imlout ;model yord=vchi ;
proc reg data=imlout ;model yca1=vchi ;
proc reg data=imlout ;model yca2=vchi ;
/*
%mend its;
%its
*/
run;
```

APPENDIX C

AFROTC FORM 708

CADET FIELD TRAINING PERFORMANCE REPORT

I. RATEE IDENTIFICATION DATA (Read AFROTCR 45-3 carefully before filling in)

1. NAME (Last, First, MI)	2. SSN	3. PERIOD OF REPORT FROM: TO:
4. FIELD TRAINING BASE/SESSION	5. DET	6. CATEGORY

7. AWARDS RECEIVED

- [ ] Commandant
- [ ] Vice-Commandant
- [ ] Superior Performance
- [ ] Athletic Leadership
- [ ] Athletic
- [ ] Fleetfoot
- [ ] Academic
- [ ] Marksmanship

II. LEADERSHIP POSITIONS AND/OR ADDITIONAL DUTIES

- [ ] GROUP COMMANDER
- [ ] GROUP EXECUTIVE OFFICER
- [ ] GROUP OPERATIONS OFFICER
- [ ] GROUP ADMINISTRATIVE OFFICER
- [ ] SQUADRON COMMANDER
- [ ] FLIGHT COMMANDER
- [ ] FLIGHT ADJUTANT
- [ ] GROUP/FLIGHT SSO
- [ ] GROUP/FLIGHT STANDARDIZATION
- [ ] OTHER

III. FACTORS

A. PERFORMANCE FACTORS

1 - DOES NOT MEET STANDARDS
2 - MEETS STANDARDS BUT NEEDS IMPROVEMENT
3 - MEETS STANDARDS
4 - EXCEEDS STANDARDS

1. ADAPTABILITY TO MILITARY TRAINING (respects authority, adheres to standards, and exercises self-discipline)	
2. DUTY PERFORMANCE (effort, judgment, and self-confidence)	
3. LEADERSHIP/FOLLOWERSHIP (decisive, displays initiative, and supports others who lead)	
4. ADAPTABILITY TO STRESS (stable, flexible, and dependable)	
5. DRILL AND CEREMONIES (command voice, precision, bearing, alignment, and execution)	
6. HUMAN RELATIONS (sensitivity, cooperation, empathy, and attitude)	
7. PHYSICAL FITNESS (timed runs and physical fitness tests)	
8. COMMUNICATION SKILLS (clear, concise, and organized)	
9. JUDGMENT AND DECISIONS (organizational skills, time management, and accepts responsibility)	
10. PROFESSIONAL QUALITIES (appearance, customs and courtesies, and bearing)	

B. INDIVIDUAL SCORING RESULTS

1. BEST 1.5 MILE RUN TIME:
2. BEST PFT SCORE:
3. AS100 AVG:
4. AS200 AVG:
5. ACADEMIC AVG:

C. OVERALL PERFORMANCE FACTORS

1. RATEE AVG:
2. FLIGHT AVG:
3. DIFFERENTIAL: (+ or -)

D. LOWER 25% [ ] YES [ ] NO

E. MEETS STANDARDS [ ] YES [ ] NO

AFROTC FORM 708, FEB 91 (COMPUTER GENERATED) PREVIOUS EDITIONS ARE OBSOLETE

IV. FTO COMMENTS	(CADET NAME: ^)
^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^	

NAME, GRADE, DET AND LOCATION	DUTY TITLE	DATE
^ ^ DET ^ ^	Field Training Officer	^
	SSN ^	SIGNATURE

V. COC COMMENTS	^	CONCUR	^	NONCONCUR

^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^

NAME, GRADE, DET AND LOCATION	DUTY TITLE	DATE
^ ^ DET ^ ^	Commandant of Cadets	^
	SSN ^	SIGNATURE

VI. FTCC COMMENTS	^	CONCUR	^	NONCONCUR

^ ^ ^ ^ ^ ^

NAME, GRADE, DET AND LOCATION	DUTY TITLE	DATE
^ ^ DET ^	Commander	^
	SSN ^	SIGNATURE

AFROTC FORM 708, FEB 91 (REVERSE)

APPENDIX D

COMPARATIVE RESULTS PREDICTING FIELD TRAINING PERFORMANCE SCORES USING THREE TYPES OF EMPIRICAL AND ONE RATIONAL KEY

LEAP Subconstructs	Rational key	Empirical Key Approaches		
		Nominal	Ordinal	Correspondence Analysis
LEAP TOT	.11	.61	.45	.31
Trf Ldr	.03	.64	.21	.19
Trn Ldr	-.05	.31	.04	.07
D-M Abl	.05	.26	.22	.07
G/S Inf	.07	.28	.21	.13
T-P Or	.10	.27	.15	.01
S-S Or	.07	.29	.25	.16
Phy Fit	.19	.37	.35	.31
Inst Com	.05	.23	.14	.08
Prs Excl	.11	.28	.25	.12
Tol Adv	.10	.20	.10	.10
Soc Pwr	.01	.33	.22	.20
Ret Prp	-.04	.18	.08	.01

APPENDIX E

PEER RATINGS FORM

AFROTC PEER RATING FORM

DIRECTIONS: Rate each of the 8 randomly selected cadets listed on your Peer Rating List using the "Almost Never" to "Almost Always" scale below to indicate the frequency of times you observed their behaviors on each of the 19 dimensions listed on the pages which follow. If you do not have enough information concerning a particular behavior for a certain cadet, please mark the "F" response.

A	B	C	D	E	F
Almost Never	**Infrequently**	**Sometimes**	**Frequently**	**Almost Always**	**Not Enough Information**

RECORD EACH OF YOUR RATINGS ON THE MACHINE SCORABLE ANSWER SHEETS. Mark your answer on the appropriate line number (as given on this form) and in the space corresponding to the A to E rating you've selected. Do this for each dimension and for each cadet rated. You will be rating all 8 cadets on a given dimension before going on to the next dimension.

YOU SHOULD NOT IDENTIFY YOURSELF ON THE ANSWER SHEET.

However, so that is possible to identify who you are rating, in the section called, "Special Codes," under Column K, enter the number of your FLIGHT as indicated on your PEER RATING LIST. Under Column L, enter the number of the GROUP to which your rated cadets belong (1 or 2), as indicated on the top of your Peer Rating List.

A----------B----------C----------D----------E F

Almost Never	Infrequently	Sometimes	Frequently	Almost Always	Not Enough Information

DIMENSION 1: When serving as the leader, this cadet motivated others to go beyond their best previous levels of performance.

1. First Cadet on your List
2. Second Cadet on your List
3. Third Cadet on your List
4. Fourth Cadet on your List
5. Fifth Cadet on your List
6. Sixth Cadet on your List
7. Seventh Cadet on your List
8. Eighth Cadet on your List

DIMENSION 2: When serving as the leader, this cadet rewarded good performance and reprimanded poor performance of others.

9. First Cadet on your List
10. Second Cadet on your List
11. Third Cadet on your List
12. Fourth Cadet on your List
13. Fifth Cadet on your List
14. Sixth Cadet on your List
15. Seventh Cadet on your List
16. Eighth Cadet on your List

DIMENSION 3: This cadet was able to identify problems, analyze them, and then come up with effective solutions.

17. First Cadet on your List
18. Second Cadet on your List
19. Third Cadet on your List
20. Fourth Cadet on your List
21. Fifth Cadet on your List
22. Sixth Cadet on your List
23. Seventh Cadet on your List
24. Eighth Cadet on your List

CONTINUE USING THE FOLLOWING SCALE TO INDICATE THE FREQUENCY OF BEHAVIOR FOR EACH CADET ON THE FOLLOWING DIMENSIONS. IF YOU DON'T HAVE ENOUGH INFORMATION ON A CADET FOR A PARTICULAR DIMENSION, MARK RESPONSE "F".

A	B	C	D	E	F
Almost Never	Infrequently	Sometimes	Frequently	Almost Always	Not Enough Information

DIMENSION 4: By monitoring what was going on, this cadet gathered useful information, then shared it with others so that it could be used to help the flight better carry out its work.

25. First Cadet on your List
26. Second Cadet on your List
27. Third Cadet on your List
28. Fourth Cadet on your List
29. Fifth Cadet on your List
30. Sixth Cadet on your List
31. Seventh Cadet on your List
32. Eighth Cadet on your List

DIMENSION 5: This cadet worked well with other flight members, drawing on each cadet's ideas, strengths or resources to collaboratively achieve the group's goals.

33. First Cadet on your List
34. Second Cadet on your List
35. Third Cadet on your List
36. Fourth Cadet on your List
37. Fifth Cadet on your List
38. Sixth Cadet on your List
39. Seventh Cadet on your List
40. Eighth Cadet on your List

CONTINUE USING THE FOLLOWING SCALE TO INDICATE THE FREQUENCY OF BEHAVIOR FOR EACH CADET ON THE FOLLOWING DIMENSIONS. IF YOU DON'T HAVE ENOUGH INFORMATION ON A CADET FOR A PARTICULAR DIMENSION, MARK RESPONSE "F".

A	B	C	D	E	F
Almost Never	**Infrequently**	**Sometimes**	**Frequently**	**Almost Always**	**Not Enough Information**

DIMENSION 6: This cadet worked effectively on his or her own, relying on his or her own judgment to make needed decisions.

41. First Cadet on your List
42. Second Cadet on your List
43. Third Cadet on your List
44. Fourth Cadet on your List
45. Fifth Cadet on your List
46. Sixth Cadet on your List
47. Seventh Cadet on your List
48. Eighth Cadet on your List

DIMENSION 7: This cadet showed a concern for maintaining good health through willing participation in more than the required physical conditioning.

49. First Cadet on your List
50. Second Cadet on your List
51. Third Cadet on your List
52. Fourth Cadet on your List
53. Fifth Cadet on your List
54. Sixth Cadet on your List
55. Seventh Cadet on your List
56. Eighth Cadet on your List

CONTINUE USING THE FOLLOWING SCALE TO INDICATE THE FREQUENCY OF BEHAVIOR FOR EACH CADET ON THE FOLLOWING DIMENSIONS. IF YOU DON'T HAVE ENOUGH INFORMATION ON A CADET FOR A PARTICULAR DIMENSION, MARK RESPONSE "F".

A	B	C	D	E	F
Almost Never	Infrequently	Sometimes	Frequently	Almost Always	Not Enough Information

DIMENSION 8: This cadet willingly made personal sacrifices out of loyalty to the Air Force or out of commitment to its goals and values.

57. First Cadet on your List
58. Second Cadet on your List
59. Third Cadet on your List
60. Fourth Cadet on your List **(TURN ANSWER SHEET TO SIDE 2 AND CONTINUE)**
61. Fifth Cadet on your List
62. Sixth Cadet on your List
63. Seventh Cadet on your List
64. Eighth Cadet on your List

DIMENSION 9: This cadet worked hard on assigned duties and tasks, and was not satisfied until the best possible performance was achieved.

65. First Cadet on your List
66. Second Cadet on your List
67. Third Cadet on your List
68. Fourth Cadet on your List
69. Fifth Cadet on your List
70. Sixth Cadet on your List
71. Seventh Cadet on your List
72. Eighth Cadet on your List

CONTINUE USING THE FOLLOWING SCALE TO INDICATE THE FREQUENCY OF BEHAVIOR FOR EACH CADET ON THE FOLLOWING DIMENSIONS. IF YOU DON'T HAVE ENOUGH INFORMATION ON A CADET FOR A PARTICULAR DIMENSION, MARK RESPONSE "F".

A	B	C	D	E	F
Almost Never	Infrequently	Sometimes	Frequently	Almost Always	Not Enough Information

DIMENSION 10. This cadet worked hard at all duties or tasks despite any adversity or frustration experienced.

73. First Cadet on your List
74. Second Cadet on your List
75. Third Cadet on your List
76. Fourth Cadet on your List
77. Fifth Cadet on your List
78. Sixth Cadet on your List
79. Seventh Cadet on your List
80. Eighth Cadet on your List

DIMENSION 11: This cadet listened to, advised and supported others.

81. First Cadet on your List
82. Second Cadet on your List
83. Third Cadet on your List
84. Fourth cadet on your List
85. Fifth Cadet on your List
86. Sixth Cadet on your List
87. Seventh Cadet on your List
88. Eighth Cadet on your List

CONTINUE USING THE FOLLOWING SCALE TO INDICATE THE FREQUENCY OF BEHAVIOR FOR EACH CADET ON THE FOLLOWING DIMENSIONS. IF YOU DON'T HAVE ENOUGH INFORMATION ON A CADET FOR A PARTICULAR DIMENSION, MARK RESPONSE "F".

A----------B----------C----------D----------E F

Almost Never	Infrequently	Sometimes	Frequently	Almost Always	Not Enough Information

DIMENSION 12. This cadet encouraged others to take the work of the flight more seriously, and to make a stronger commitment to the achievement of its goals.

89. First Cadet on your List
90. Second Cadet on your List
91. Third Cadet on your List
92. Fourth Cadet on your List
93. Fifth Cadet on your List
94. Sixth Cadet on your List
95. Seventh Cadet on your List
96. Eighth Cadet on your List

DIMENSION 13: This cadet inspired others and gained their support for his/her suggestions and ideas.

97. First Cadet on your List
98. Second Cadet on your List
99. Third Cadet on your List
100. Fourth Cadet on your List
101. Fifth Cadet on your List
102. Sixth Cadet on your List
103. Seventh Cadet on your List
104. Eighth Cadet on your List

CONTINUE USING THE FOLLOWING SCALE TO INDICATE THE FREQUENCY OF BEHAVIOR FOR EACH CADET ON THE FOLLOWING DIMENSIONS. IF YOU DON'T HAVE ENOUGH INFORMATION ON A CADET FOR A PARTICULAR DIMENSION, MARK RESPONSE "F".

A	B	C	D	E	F
Almost Never	**Infrequently**	**Sometimes**	**Frequently**	**Almost Always**	**Not Enough Information**

DIMENSION 14: This cadet found new and creative ways to solve problems or complete tasks.

105. First Cadet on your List
106. Second Cadet on your List
107. Third Cadet on your List
108. Fourth Cadet on your List
109. Fifth Cadet on your List
110. Sixth Cadet on your List
111. Seventh Cadet on your List
112. Eighth Cadet on your List

DIMENSION 15: In a leadership position, he/she considered the needs and abilities of others when assigning tasks or duties.

113. First Cadet on your List
114. Second Cadet on your List
115. Third Cadet on your List
116. Fourth Cadet on your List
117. Fifth Cadet on your List
118. Sixth Cadet on your List
119. Seventh Cadet on your List
120. Eighth Cadet on your List

FOR DIMENSIONS 16 - 19, BEGIN ON ANSWER SHEET 2 SIDE 1

CONTINUE USING THE FOLLOWING SCALE TO INDICATE THE FREQUENCY OF BEHAVIOR FOR EACH CADET ON THE FOLLOWING DIMENSIONS. IF YOU DON'T HAVE ENOUGH INFORMATION ON A CADET FOR A PARTICULAR DIMENSION, MARK RESPONSE "F".

A----------B----------C----------D----------E F

Almost Never	Infrequently	Sometimes	Frequently	Almost Always	Not Enough Information

DIMENSION 16: This cadet motivated others to act by raising challenging problems or questions for them to solve. This cadet helped others find new ways to think and to handle tasks or assignments.

1. First Cadet on your List
2. Second Cadet on your List
3. Third Cadet on your List
4. Fourth Cadet on your List
5. Fifth Cadet on your List
6. Sixth Cadet on your List
7. Seventh Cadet on your List
8. Eighth Cadet on your List

DIMENSION 17: This cadet planned and carried out activities in an organized fashion.

9. First Cadet on your List
10. Second Cadet on your List
11. Third Cadet on your List
12. Fourth Cadet on your List
13. Fifth Cadet on your List
14. Sixth Cadet on your List
15. Seventh Cadet on your List
16. Eighth Cadet on your List

CONTINUE USING THE FOLLOWING SCALE TO INDICATE THE FREQUENCY OF BEHAVIOR FOR EACH CADET ON THE FOLLOWING DIMENSIONS. IF YOU DON'T HAVE ENOUGH INFORMATION ON A CADET FOR A PARTICULAR DIMENSION, MARK RESPONSE "F".

A----------B----------C----------D----------E F

Almost Never	Infrequently	Sometimes	Frequently	Almost Always	Not Enough Information

DIMENSION 18: This cadet demonstrated qualities that resulted in a high degree of success during this encampment.

17. First Cadet on your List
18. Second Cadet on your List
19. Third Cadet on your List
20. Fourth Cadet on your List
21. Fifth Cadet on your List
22. Sixth Cadet on your List
23. Seventh Cadet on your List
24. Eighth Cadet on your List

DIMENSION 19: This cadet demonstrated qualities that show the potential to be an outstanding future Air Force officer.

25. First Cadet on your List
26. Second Cadet on your List
27. Third Cadet on your List
28. Fourth Cadet on your List
29. Fifth Cadet on your List
30. Sixth Cadet on your List
31. Seventh Cadet on your List
32. Eighth Cadet on your List

APPENDIX F

VALIDATION OF THE LEAD O-2D (ROTC) SCALES AGAINST CORRESPONDING PEERING RATINGS DIMENSIONS, BASED ON A RATIONAL AND AN ORDINAL KEY

Table F-1. Validation of the LEAP O-2D (ROTC) Scales Against Corresponding Peer Rating Dimensions, Based on the Ordinal Key[a]

	19 Peer Rating Dimensions																							
LEAP Scales	**#1[b]**	**#2**	**#3**	**#4**	**#5**	**#6**	**#7**	**#8**	**#9**	**#10**	**#11**	**#12**	**#13**	**#14**	**#15**	**#16**	**#17**	**TOT DIM**	**#18**	**#19**	**TOT 18/ 19**	**TOT Trf Ldr**	**TOT D-M Abl**	**TOT Soc Pwr**
TotLEAP[c]	.11	.10	.28	.36	.21	.23	.17	.25	.18	.24	.35	.14	.09	.18	.07	.07	.26	.22	.25	.25	.27	.10	.27	.29
TrfLdr	.03	-.02	.12	.18	.04	.10	.06	.04	.04	.05	.08	.03	-.08	-.01	-.06	.01	.14	.06	.06	.13	.10	-.02	.11	.07
Chrs	.05	.02	.14	.19	.06	.00	.04	.09	.03	.15	.17	.01	-.08	.05	-.03	-.01	.11	.06	.14	.05	.11	-.01	.13	.10
TrnLdr	.08	.10	.00	.05	.02	.11	-.02	.05	.03	.04	.13	.21	.14	.16	.11	.12	.09	.13	-.07	.17	.05	.12	.07	.19
D-MAbl	-.18	-.04	-.06	-.01	.08	.18	.01	-.02	-.09	.01	.03	-.26	-.04	-.02	-.07	-.15	-.03	-.02	.06	.02	.05	-.12	-.04	-.13
G/SInf	.10	.12	.22	.26	.21	.28	.25	.27	.12	.26	.29	.13	.10	.20	-.03	.08	.28	.24	.19	.22	.22	.08	.25	.25
T-POr	-.04	-.09	.05	.03	.13	.04	.05	.14	-.01	-.04	.00	.05	.00	-.01	-.01	-.01	.02	-.01	.13	.09	.12	.00	.03	.03
S-SOr	.14	.11	.24	.26	.28	.36	.25	.27	.19	.19	.28	.07	.25	.29	.18	.17	.24	.30	.23	.22	.25	.20	.27	.21
PhyFit	.17	.06	.22	.20	.09	.09	.10	.17	.20	.24	.46	.13	.04	.14	.16	.04	.18	.15	.19	.14	.18	.11	.20	.35
InstCom	.22	.11	.23	.22	.09	-.01	.06	.15	.16	.11	.05	.21	.06	.18	.13	.05	.14	.11	.17	.08	.14	.12	.19	.14
PrsExcl	.00	.20	.13	.17	.15	.02	.03	.05	.09	.19	.17	.10	.13	.14	.01	.05	.12	.08	.13	.08	.11	.06	.16	.15
TolAdv	-.09	-.12	-.06	.00	-.11	-.09	-.02	-.01	.00	-.02	.10	-.11	-.04	-.17	-.01	.05	-.13	-.05	-.02	-.03	-.03	-.02	-.11	.00
SocPwr	.09	.08	.11	.18	-.02	.02	-.02	.13	.10	.12	.20	.13	.02	.02	.03	-.03	.10	.06	.05	.02	.04	.02	.09	.19

[a] n = 59

[b] #1 = motivates others to go beyond previous best
#2 = rewards good, reprimands poor performance
#3 = identifies problems, determines solutions
#4 = gathers information, shares with others
#5 = works well with others, draws on their talents
#6 = works well on own, relies on own judgment
#7 = maintains good health through physical training
#8 = makes personal sacrifices for the Air Force
#9 = works hard until best performance is achieved
#10 = works hard despite adversities
#11 = listens to, advises, and supports others
#12 = encourages others to take work seriously
#13 = inspires others and gains their support
#14 = finds new ways to solve problems
#15 = assigns duties based on needs and abilities
#16 = motivates others by presenting challenging tasks
#17 = plans/carries out tasks in organized manner
TOT DIM = composite of dimensions 1-17
#18 = demonstrates successful qualities in present
#19 = shows qualities for success in future
TOT 18/19 = composite of dimensions #18 and #19
TOT TrfLdr = composite of #1, #13, #15 and #16
TOT D-MAbl = composite of #14 and #17
TOT SocPwr = composite of #11 and #12

[c] TotLEAP = Total LEAP Score
TrfLdr = Transformational Leadership
Chrs = Charisma
TrnLdr = Transactional Leadership
D-MAbl = Decision-Making Abilities
G/SInf = Giving/Seeking Information
T-POr = Team Player Orientation
S-SOr = Self-Sufficiency Orientation
PhyFit = Physical Fitness Factors
InstCom = Institutional Commitment
PrsExcl = Persistence to Excellence
TolAdv = Toleration of Adversity
SocPwr = Socialized Power

Table F-2. Validation of the LEAP O-2D (ROTC) Scales Against Corresponding Peer Rating, Based on the Rational Key[a]

	19 Peer Rating Dimensions																								
LEAP Scales	**#1[b]**	**#2**	**#3**	**#4**	**#5**	**#6**	**#7**	**#8**	**#9**	**#10**	**#11**	**#12**	**#13**	**#14**	**#15**	**#16**	**#17**	**TOT DIM**	**#18**	**#19**	**TOT 18/ 19**	**TOT Trf Ldr**	**TOT D-M Abl**	**TOT Soc Pwr**
TotLEAP[c]	.21	.16	.19	.16	.12	.18	.13	.15	.18	.18	.10	.15	.12	.16	.05	.11	.13	.21	.17	.15	.17	.16	.18	.14
TrfLdr	.23	.15	.20	.18	.15	.17	.12	.13	.13	.17	.15	.17	.15	.10	.10	.13	.10	.21	.20	.21	.21	.20	.16	.18
Chrs	.15	.11	.18	.17	.14	.10	.10	.10	.08	.16	.12	.12	.15	.11	.06	.14	.04	.17	.13	.14	.14	.17	.13	.14
TrnLdr	-.06	.07	-.06	-.07	-.03	-.07	-.05	-.05	-.04	-.02	-.05	-.08	-.05	.07	-.07	-.08	.00	-.04	.02	.01	.00	-.08	.00	-.07
D-MAbl	.12	.17	.10	.12	.12	.13	.25	.11	.11	.09	.05	.10	.06	.05	.18	.14	.10	.18	.15	.19	.17	.15	.10	.09
G/SInf	.16	.13	.24	.26	.22	.18	.16	.19	.18	.29	.35	.14	.16	.21	.05	.07	.33	.25	.19	.18	.19	.13	.30	.29
T-POr	.20	.07	.09	.07	.07	.10	.14	.17	.16	.08	.03	.12	.02	.08	.04	.00	.09	.11	.14	.07	.12	.07	.10	.08
S-SOr	-.06	-.01	.08	.08	.03	.02	.01	.04	.01	.04	.07	-.05	.00	.09	.05	.00	.04	.03	-.03	.00	-.02	-.03	.09	.02
PhyFit	.25	.22	.26	.26	.27	.22	.19	.19	.25	.34	.24	.22	.21	.28	.16	.22	.31	.31	.29	.25	.29	.26	.34	.27
InstCom	.01	-.11	.03	-.06	-.13	.00	-.11	.10	.04	.05	-.14	.05	-.17	-.09	-.19	-.11	.00	-.08	-.06	.07	-.02	-.13	-.02	-.06
PrsExcl	.05	.22	.10	.16	.16	.00	.10	.02	.06	.16	.05	.02	.20	.19	.10	.21	.14	.16	.06	.07	.06	.18	.17	.04
TolAdv	.06	.01	.03	-.04	.06	.00	.02	-.06	.07	-.04	-.03	-.03	.06	.08	-.06	.05	-.07	.01	-.07	-.08	-.08	.03	.01	-.04
SocPwr	.05	.10	-.01	-.01	-.05	.13	.06	.05	.12	.04	-.08	.12	.03	.00	.02	.03	-.07	.06	-.01	.02	.00	.04	-.04	.02

[a]n = 161

[b]#1 = motivates others to go beyond previous best
#2 = rewards good, reprimands poor performance
#3 = identifies problems, determines solutions
#4 = gathers information, shares with others
#5 = works well with others, draws on their talents
#6 = works well on own, relies on own judgment
#7 = maintains good health through physical training
#8 = makes personal sacrifices for the Air Force
#9 = works hard until best performance is achieved
#10 = works hard despite adversities
#11 = listens to, advises, and supports others
#12 = encourages others to take work seriously
#13 = inspires others and gains their support
#14 = finds new ways to solve problems
#15 = assigns duties based on needs and abilities
#16 = motivates others by presenting challenging tasks
#17 = plans/carries out tasks in organized manner
TOT DIM = composite of dimensions 1-17
#18 = demonstrates successful qualities in present
#19 = shows qualties for success in future
TOT 18/19 = composite of dimensions #18 and #19
TOT TrfLdr = composite of #1, #13, #15 and #16
TOT D-MAbl = composite of #14 and #17
TOT SocPwr = composite of #11 and #12

[c]TotLEAP = Total LEAP Score
TrfLdr = Transformational Leadership
Chrs = Charisma
TrnLdr = Transactional Leadership
D-MAbl = Decision-Making Abilities
G/SInf = Giving/Seeking Information
T-POr = Team Player Orientation
S-SOr = Self-Sufficiency Orientation
PhyFit = Physical Fitness Factors
InstCom = Institutional Commitment
PrsExcl = Persistence to Excellence
TolAdv = Toleration of Adversity
SocPwr = Socialized Power

*U.S. GOVERNMENT PRINTING OFFICE 1993-0-701-052 60058